

6	CGENKDYDCGSKYECDDKKCKYGVBEREDDEPNVPCLVRVCHQDVCVEGGFYRNKDCKV	65
	:   :	
24	CGFNRVVKSCGT-GCPETC-----ENPDPC-DRACHQGCFCSKGLLQIDIGNCI	71
	:   :	
66	SAEDC	70
72	SPDKC	76









QY 58 RNKDKCVSAEDCE 71  
Db 3464 LN-SDKCVLRAECD 3476

RESULT 5  
LRP2\_RAT STANDARD; PRT; 4660 AA.  
AC P8158;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalyn)  
DE (Glycoprotein 330) (gp330).  
GN LRP2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;  
RX MEDLINE=95024033; PubMed=7937880;  
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;  
RT "Complete cloning and sequencing of rat gp330/'megalyn,' a  
RT distinctive member of the low density lipoprotein receptor gene  
RT family.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).  
RN [2]  
RP FUNCTION.  
RX MEDLINE=9538696; PubMed=7544804;  
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,  
RA Norris K., Gliemann J., Christensen E.I.;  
RT "Evidence that epithelial glycoprotein 330/megalyn mediates uptake of  
RT polybasic drugs.";  
RL J. Clin. Invest. 96:1404-1413(1995).  
RN [3]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=94172242; PubMed=7510321;  
RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,  
RA Andres G., McCluskey R.T.;  
RT "Organ distribution in rats of two members of the low-density  
RT lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the  
RT receptor-associated protein (RAP).";  
RL J. Histochem. Cytochem. 42:531-542(1994).  
CC -!- FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,  
CC PLASMINOGEN ACTIVATOR-PLASMINOGEN ACTIVATOR INHIBITOR TYPE I  
CC COMPLEX, APOLIPOPROTEIN E-ENRICHED BETA-VLDL, LIPOPROTEIN LIPASE,  
CC LACTOFERRIN, CLUSTERIN AND CALCIUM.  
CC -!- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS  
CC APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B.  
CC -!- SUBUNIT: Forms a multimeric complex together with a receptor-  
CC associated protein (RAP). Binds to ankyrin-repeat family A protein  
CC 2 (ANKRA2) (By similarity).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXPRESSED IN  
CC CLATHRIN-COATED PITS; A SOLUBLE FORM IS POSSIBLY DERIVED BY  
CC CLEAVAGE AT THE CELL SURFACE.  
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF KIDNEY GLOMERULUS AND  
CC PROXIMAL TUBULE, LUNG, EPIDIDYMIS, YOLK SAC, AMONG OTHERS.  
CC -!- SIMILARITY: Contains 36 LDL-receptor class A domains.  
CC -!- SIMILARITY: Contains 37 LDL-receptor class B domains.  
CC -!- SIMILARITY: Contains 17 EGF-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L34049; AAA51369.1; -

PIR; T42737; T42737.  
DR HSP; Q07954; 1CR8.  
DR GlycoSuiteDB; P98158.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; Ldl\_receptor\_rep.  
DR Pfam; PF00008; EGF\_13.  
DR Pfam; PF00057; ldl\_recept\_a; 36.  
DR Pfam; PF00058; ldl\_recept\_b; 33.  
DR SMART; SM00179; EGF\_CA; 3.  
DR SMART; SM00192; LDLA; 36.  
DR SMART; SM00135; LY; 34.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 8.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLRA\_1; 31.  
DR PROSITE; PS00068; LDLRA\_2; 36.  
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;  
KW Receptor; EGF-like domain; SH3-binding; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 4660  
FT DOMAIN 26 4425  
FT TRANSMEM 4426 4446  
FT DOMAIN 4447 4660  
FT DOMAIN 26 64  
FT DOMAIN 65 105  
FT DOMAIN 106 144  
FT DOMAIN 145 181  
FT DOMAIN 181 219  
FT DOMAIN 220 258  
FT DOMAIN 283 307  
FT DOMAIN 308 346  
FT DOMAIN 347 385  
FT DOMAIN 435 476  
FT DOMAIN 478 519  
FT DOMAIN 521 566  
FT DOMAIN 568 611  
FT DOMAIN 612 652  
FT DOMAIN 658 704  
FT DOMAIN 752 793  
FT DOMAIN 795 835  
FT DOMAIN 837 879  
FT DOMAIN 881 923  
FT DOMAIN 969 1013  
FT DOMAIN 1023 1061  
FT DOMAIN 1064 1103  
FT DOMAIN 1108 1146  
FT DOMAIN 1148 1186  
FT DOMAIN 1186 1225  
FT DOMAIN 1229 1269  
FT DOMAIN 1270 1308  
FT DOMAIN 1311 1351  
FT DOMAIN 1350 1390  
FT DOMAIN 1391 1430  
FT DOMAIN 1479 1520  
FT DOMAIN 1522 1563  
FT DOMAIN 1567 1609  
FT DOMAIN 1611 1654  
FT DOMAIN 1656 1696  
FT DOMAIN 1701 1742  
FT DOMAIN 1791 1832  
FT DOMAIN 1834 1882  
FT DOMAIN 1884 1930  
FT DOMAIN 1932 1972  
FT DOMAIN 1973 2013  
FT DOMAIN 2019 2060  
FT DOMAIN 2108 2156  
FT DOMAIN 2158 2201  
FT DOMAIN 2203 2245  
FT DOMAIN 2245 2285  
FT DOMAIN 2285 2325  
FT DOMAIN 2325 2365  
FT DOMAIN 2365 2405  
FT DOMAIN 2405 2445  
FT DOMAIN 2445 2485  
FT DOMAIN 2485 2525  
FT DOMAIN 2525 2565  
FT DOMAIN 2565 2605  
FT DOMAIN 2605 2645  
FT DOMAIN 2645 2685  
FT DOMAIN 2685 2725  
FT DOMAIN 2725 2765  
FT DOMAIN 2765 2805  
FT DOMAIN 2805 2845  
FT DOMAIN 2845 2885  
FT DOMAIN 2885 2925  
FT DOMAIN 2925 2965  
FT DOMAIN 2965 3005  
FT DOMAIN 3005 3045  
FT DOMAIN 3045 3085  
FT DOMAIN 3085 3125  
FT DOMAIN 3125 3165  
FT DOMAIN 3165 3205  
FT DOMAIN 3205 3245  
FT DOMAIN 3245 3285  
FT DOMAIN 3285 3325  
FT DOMAIN 3325 3365  
FT DOMAIN 3365 3405  
FT DOMAIN 3405 3445  
FT DOMAIN 3445 3485  
FT DOMAIN 3485 3525  
FT DOMAIN 3525 3565  
FT DOMAIN 3565 3605  
FT DOMAIN 3605 3645  
FT DOMAIN 3645 3685  
FT DOMAIN 3685 3725  
FT DOMAIN 3725 3765  
FT DOMAIN 3765 3805  
FT DOMAIN 3805 3845  
FT DOMAIN 3845 3885  
FT DOMAIN 3885 3925  
FT DOMAIN 3925 3965  
FT DOMAIN 3965 4005  
FT DOMAIN 4005 4045  
FT DOMAIN 4045 4085  
FT DOMAIN 4085 4125  
FT DOMAIN 4125 4165  
FT DOMAIN 4165 4205  
FT DOMAIN 4205 4245  
FT DOMAIN 4245 4285  
FT DOMAIN 4285 4325  
FT DOMAIN 4325 4365  
FT DOMAIN 4365 4405  
FT DOMAIN 4405 4446  
FT DOMAIN 4446 4486  
FT DOMAIN 4486 4526  
FT DOMAIN 4526 4566  
FT DOMAIN 4566 4606  
FT DOMAIN 4606 4646  
FT DOMAIN 4646 4686  
FT DOMAIN 4686 4726  
FT DOMAIN 4726 4766  
FT DOMAIN 4766 4806  
FT DOMAIN 4806 4846  
FT DOMAIN 4846 4886  
FT DOMAIN 4886 4926  
FT DOMAIN 4926 4966  
FT DOMAIN 4966 5006  
FT DOMAIN 5006 5046  
FT DOMAIN 5046 5086  
FT DOMAIN 5086 5126  
FT DOMAIN 5126 5166  
FT DOMAIN 5166 5206  
FT DOMAIN 5206 5246  
FT DOMAIN 5246 5286  
FT DOMAIN 5286 5326  
FT DOMAIN 5326 5366  
FT DOMAIN 5366 5406  
FT DOMAIN 5406 5446  
FT DOMAIN 5446 5486  
FT DOMAIN 5486 5526  
FT DOMAIN 5526 5566  
FT DOMAIN 5566 5606  
FT DOMAIN 5606 5646  
FT DOMAIN 5646 5686  
FT DOMAIN 5686 5726  
FT DOMAIN 5726 5766  
FT DOMAIN 5766 5806  
FT DOMAIN 5806 5846  
FT DOMAIN 5846 5886  
FT DOMAIN 5886 5926  
FT DOMAIN 5926 5966  
FT DOMAIN 5966 6006  
FT DOMAIN 6006 6046  
FT DOMAIN 6046 6086  
FT DOMAIN 6086 6126  
FT DOMAIN 6126 6166  
FT DOMAIN 6166 6206  
FT DOMAIN 6206 6246  
FT DOMAIN 6246 6286  
FT DOMAIN 6286 6326  
FT DOMAIN 6326 6366  
FT DOMAIN 6366 6406  
FT DOMAIN 6406 6446  
FT DOMAIN 6446 6486  
FT DOMAIN 6486 6526  
FT DOMAIN 6526 6566  
FT DOMAIN 6566 6606  
FT DOMAIN 6606 6646  
FT DOMAIN 6646 6686  
FT DOMAIN 6686 6726  
FT DOMAIN 6726 6766  
FT DOMAIN 6766 6806  
FT DOMAIN 6806 6846  
FT DOMAIN 6846 6886  
FT DOMAIN 6886 6926  
FT DOMAIN 6926 6966  
FT DOMAIN 6966 7006  
FT DOMAIN 7006 7046  
FT DOMAIN 7046 7086  
FT DOMAIN 7086 7126  
FT DOMAIN 7126 7166  
FT DOMAIN 7166 7206  
FT DOMAIN 7206 7246  
FT DOMAIN 7246 7286  
FT DOMAIN 7286 7326  
FT DOMAIN 7326 7366  
FT DOMAIN 7366 7406  
FT DOMAIN 7406 7446  
FT DOMAIN 7446 7486  
FT DOMAIN 7486 7526  
FT DOMAIN 7526 7566  
FT DOMAIN 7566 7606  
FT DOMAIN 7606 7646  
FT DOMAIN 7646 7686  
FT DOMAIN 7686 7726  
FT DOMAIN 7726 7766  
FT DOMAIN 7766 7806  
FT DOMAIN 7806 7846  
FT DOMAIN 7846 7886  
FT DOMAIN 7886 7926  
FT DOMAIN 7926 7966  
FT DOMAIN 7966 8006  
FT DOMAIN 8006 8046  
FT DOMAIN 8046 8086  
FT DOMAIN 8086 8126  
FT DOMAIN 8126 8166  
FT DOMAIN 8166 8206  
FT DOMAIN 8206 8246  
FT DOMAIN 8246 8286  
FT DOMAIN 8286 8326  
FT DOMAIN 8326 8366  
FT DOMAIN 8366 8406  
FT DOMAIN 8406 8446  
FT DOMAIN 8446 8486  
FT DOMAIN 8486 8526  
FT DOMAIN 8526 8566  
FT DOMAIN 8566 8606  
FT DOMAIN 8606 8646  
FT DOMAIN 8646 8686  
FT DOMAIN 8686 8726  
FT DOMAIN 8726 8766  
FT DOMAIN 8766 8806  
FT DOMAIN 8806 8846  
FT DOMAIN 8846 8886  
FT DOMAIN 8886 8926  
FT DOMAIN 8926 8966  
FT DOMAIN 8966 9006  
FT DOMAIN 9006 9046  
FT DOMAIN 9046 9086  
FT DOMAIN 9086 9126  
FT DOMAIN 9126 9166  
FT DOMAIN 9166 9206  
FT DOMAIN 9206 9246  
FT DOMAIN 9246 9286  
FT DOMAIN 9286 9326  
FT DOMAIN 9326 9366  
FT DOMAIN 9366 9406  
FT DOMAIN 9406 9446  
FT DOMAIN 9446 9486  
FT DOMAIN 9486 9526  
FT DOMAIN 9526 9566  
FT DOMAIN 9566 9606  
FT DOMAIN 9606 9646  
FT DOMAIN 9646 9686  
FT DOMAIN 9686 9726  
FT DOMAIN 9726 9766  
FT DOMAIN 9766 9806  
FT DOMAIN 9806 9846  
FT DOMAIN 9846 9886  
FT DOMAIN 9886 9926  
FT DOMAIN 9926 9966  
FT DOMAIN 9966 10006  
FT DOMAIN 10006 10046  
FT DOMAIN 10046 10086  
FT DOMAIN 10086 10126  
FT DOMAIN 10126 10166  
FT DOMAIN 10166 10206  
FT DOMAIN 10206 10246  
FT DOMAIN 10246 10286  
FT DOMAIN 10286 10326  
FT DOMAIN 10326 10366  
FT DOMAIN 10366 10406  
FT DOMAIN 10406 10446  
FT DOMAIN 10446 10486  
FT DOMAIN 10486 10526  
FT DOMAIN 10526 10566  
FT DOMAIN 10566 10606  
FT DOMAIN 10606 10646  
FT DOMAIN 10646 10686  
FT DOMAIN 10686 10726  
FT DOMAIN 10726 10766  
FT DOMAIN 10766 10806  
FT DOMAIN 10806 10846  
FT DOMAIN 10846 10886  
FT DOMAIN 10886 10926  
FT DOMAIN 10926 10966  
FT DOMAIN 10966 11006  
FT DOMAIN 11006 11046  
FT DOMAIN 11046 11086  
FT DOMAIN 11086 11126  
FT DOMAIN 11126 11166  
FT DOMAIN 11166 11206  
FT DOMAIN 11206 11246  
FT DOMAIN 11246 11286  
FT DOMAIN 11286 11326  
FT DOMAIN 11326 11366  
FT DOMAIN 11366 11406  
FT DOMAIN 11406 11446  
FT DOMAIN 11446 11486  
FT DOMAIN 11486 11526  
FT DOMAIN 11526 11566  
FT DOMAIN 11566 11606  
FT DOMAIN 11606 11646  
FT DOMAIN 11646 11686  
FT DOMAIN 11686 11726  
FT DOMAIN 11726 11766  
FT DOMAIN 11766 11806  
FT DOMAIN 11806 11846  
FT DOMAIN 11846 11886  
FT DOMAIN 11886 11926  
FT DOMAIN 11926 11966  
FT DOMAIN 11966 12006  
FT DOMAIN 12006 12046  
FT DOMAIN 12046 12086  
FT DOMAIN 12086 12126  
FT DOMAIN 12126 12166  
FT DOMAIN 12166 12206  
FT DOMAIN 12206 12246  
FT DOMAIN 12246 12286  
FT DOMAIN 12286 12326  
FT DOMAIN 12326 12366  
FT DOMAIN 12366 12406  
FT DOMAIN 12406 12446  
FT DOMAIN 12446 12486  
FT DOMAIN 12486 12526  
FT DOMAIN 12526 12566  
FT DOMAIN 12566 12606  
FT DOMAIN 12606 12646  
FT DOMAIN 12646 12686  
FT DOMAIN 12686 12726  
FT DOMAIN 12726 12766  
FT DOMAIN 12766 12806  
FT DOMAIN 12806 12846  
FT DOMAIN 12846 12886  
FT DOMAIN 12886 12926  
FT DOMAIN 12926 12966  
FT DOMAIN 12966 13006  
FT DOMAIN 13006 13046  
FT DOMAIN 13046 13086  
FT DOMAIN 13086 13126  
FT DOMAIN 13126 13166  
FT DOMAIN 13166 13206  
FT DOMAIN 13206 13246  
FT DOMAIN 13246 13286  
FT DOMAIN 13286 13326  
FT DOMAIN 13326 13366  
FT DOMAIN 13366 13406  
FT DOMAIN 13406 13446  
FT DOMAIN 13446 13486  
FT DOMAIN 13486 13526  
FT DOMAIN 13526 13566  
FT DOMAIN 13566 13606  
FT DOMAIN 13606 13646  
FT DOMAIN 13646 13686  
FT DOMAIN 13686 13726  
FT DOMAIN 13726 13766  
FT DOMAIN 13766 13806  
FT DOMAIN 13806 13846  
FT DOMAIN 13846 13886  
FT DOMAIN 13886 13926  
FT DOMAIN 13926 13966  
FT DOMAIN 13966 14006  
FT DOMAIN 14006 14046  
FT DOMAIN 14046 14086  
FT DOMAIN 14086 14126  
FT DOMAIN 14126 14166  
FT DOMAIN 14166 14206  
FT DOMAIN 14206 14246  
FT DOMAIN 14246 14286  
FT DOMAIN 14286 14326  
FT DOMAIN 14326 14366  
FT DOMAIN 14366 14406  
FT DOMAIN 14406 14446  
FT DOMAIN 14446 14486  
FT DOMAIN 14486 14526  
FT DOMAIN 14526 14566  
FT DOMAIN 14566 14606  
FT DOMAIN 14606 14646  
FT DOMAIN 14646 14686  
FT DOMAIN 14686 14726  
FT DOMAIN 14726 14766  
FT DOMAIN 14766 14806  
FT DOMAIN 14806 14846  
FT DOMAIN 14846 14886  
FT DOMAIN 14886 14926  
FT DOMAIN 14926 14966  
FT DOMAIN 14966 15006  
FT DOMAIN 15006 15046  
FT DOMAIN 15046 15086  
FT DOMAIN 15086 15126  
FT DOMAIN 15126 15166  
FT DOMAIN 15166 15206  
FT DOMAIN 15206 15246  
FT DOMAIN 15246 15286  
FT DOMAIN 15286 15326  
FT DOMAIN 15326 15366  
FT DOMAIN 15366 15406  
FT DOMAIN 15406 15446  
FT DOMAIN 15446 15486  
FT DOMAIN 15486 15526  
FT DOMAIN 15526 15566  
FT DOMAIN 15566 15606  
FT DOMAIN 15606 15646  
FT DOMAIN 15646 15686  
FT DOMAIN 15686 15726  
FT DOMAIN 15726 15766  
FT DOMAIN 15766 15806  
FT DOMAIN 15806 15846  
FT DOMAIN 15846 15886  
FT DOMAIN 15886 15926  
FT DOMAIN 15926 15966  
FT DOMAIN 15966 16006  
FT DOMAIN 16006 16046  
FT DOMAIN 16046 16086  
FT DOMAIN 16086 16126  
FT DOMAIN 16126 16166  
FT DOMAIN 16166 16206  
FT DOMAIN 16206 16246  
FT DOMAIN 16246 16286  
FT DOMAIN 16286 16326  
FT DOMAIN 16326 16366  
FT DOMAIN 16366 16406  
FT DOMAIN 16406 16446  
FT DOMAIN 16446 16486  
FT DOMAIN 16486 16526  
FT DOMAIN 16526 16566  
FT DOMAIN 16566 16606  
FT DOMAIN 16606 16646  
FT DOMAIN 16646 16686  
FT DOMAIN 16686 16726  
FT DOMAIN 16726 16766  
FT DOMAIN 16766 16806  
FT DOMAIN 16806 16846  
FT DOMAIN 16846 16886  
FT DOMAIN 16886 16926  
FT DOMAIN 16926 16966  
FT DOMAIN 16966 17006  
FT DOMAIN 17006 17046  
FT DOMAIN 17046 17086  
FT DOMAIN 17086 17126  
FT DOMAIN 17126 17166  
FT DOMAIN 17166 17206  
FT DOMAIN 17206 17246  
FT DOMAIN 17246 17286  
FT DOMAIN 17286 17326  
FT DOMAIN 17326 17366  
FT DOMAIN 17366 17406  
FT DOMAIN 17406 17446  
FT DOMAIN 17446 17486  
FT DOMAIN 17486 17526  
FT DOMAIN 17526 17566  
FT DOMAIN 17566 17606  
FT DOMAIN 17606 17646  
FT DOMAIN 17646 17686  
FT DOMAIN 17686 17726  
FT DOMAIN 17726 17766  
FT DOMAIN 17766 17806  
FT DOMAIN 17806 17846  
FT DOMAIN 17846 17886  
FT DOMAIN 17886 17926  
FT DOMAIN 17926 17966  
FT DOMAIN 17966 18006  
FT DOMAIN 18006 18046  
FT DOMAIN 18046 18086  
FT DOMAIN 18086 18126  
FT DOMAIN 18126 18166  
FT DOMAIN 18166 18206  
FT DOMAIN 18206 18246  
FT DOMAIN 18246 18286  
FT DOMAIN 18286 18326  
FT DOMAIN 18326 18366  
FT DOMAIN 18366 18406  
FT DOMAIN 18406 18446  
FT DOMAIN 18446 18486  
FT DOMAIN 18486 18526  
FT DOMAIN 18526 18566  
FT DOMAIN 18566 18606  
FT DOMAIN 18606 18646  
FT DOMAIN 18646 18686  
FT DOMAIN 18686 18726  
FT DOMAIN 18726 18766  
FT DOMAIN 18766 18806  
FT DOMAIN 18806 18846  
FT DOMAIN 18846 18886  
FT DOMAIN 18886 18926  
FT DOMAIN 18926 18966  
FT DOMAIN 18966 19006  
FT DOMAIN 19006 19046  
FT DOMAIN 19046 19086  
FT DOMAIN 19086 19126  
FT DOMAIN 19126 19166  
FT DOMAIN 19166 19206  
FT DOMAIN 19206 19246  
FT DOMAIN 19246 19286  
FT DOMAIN 19286 19326  
FT DOMAIN 19326 19366  
FT DOMAIN 19366 19406  
FT DOMAIN 19406 19446  
FT DOMAIN 19446 19486  
FT DOMAIN 19486 19526  
FT DOMAIN 19526 19566  
FT DOMAIN 19566 19606  
FT DOMAIN 19606 19646  
FT DOMAIN 19646 19686  
FT DOMAIN 19686 19726  
FT DOMAIN 19726 19766  
FT DOMAIN 19766 19806  
FT DOMAIN 19806 19846  
FT DOMAIN 19846 19886  
FT DOMAIN 19886 19926  
FT DOMAIN 19926 19966  
FT DOMAIN 19966 20006  
FT DOMAIN 20006 20046  
FT DOMAIN 20046 20086  
FT DOMAIN 20086 20126  
FT DOMAIN 20126 20166  
FT DOMAIN 20166 20206  
FT DOMAIN 20206 20246  
FT DOMAIN 20246 20286  
FT DOMAIN 20286 20326  
FT DOMAIN 20326 20366  
FT DOMAIN 20366 20406  
FT DOMAIN 20406 20446  
FT DOMAIN 20446 20486  
FT DOMAIN 20486 20526  
FT DOMAIN 20526 20566  
FT DOMAIN 20566 20606  
FT DOMAIN 20606 20646  
FT DOMAIN 20646 20686  
FT DOMAIN 20686 20726  
FT DOMAIN 20726 20766  
FT DOMAIN 20766 20806  
FT DOMAIN 20806 20846  
FT DOMAIN 20846 20886  
FT DOMAIN 20886 20926  
FT DOMAIN 20926 20966  
FT DOMAIN 20966 21006  
FT DOMAIN 21006 21046  
FT DOMAIN 21046 21086  
FT DOMAIN 21086 21126  
FT DOMAIN 21126 21166  
FT DOMAIN 21166 21206  
FT DOMAIN 21206 21246  
FT DOMAIN 21246 21286  
FT DOMAIN 21286 21326  
FT DOMAIN 21326 21366  
FT DOMAIN 21366 21406  
FT DOMAIN 21406 21446  
FT DOMAIN 21446 21486  
FT DOMAIN 21486 21526  
FT DOMAIN 21526 21566  
FT DOMAIN 21566 21606  
FT DOMAIN 21606 21646  
FT DOMAIN 21646 21686  
FT DOMAIN 21686 21726  
FT DOMAIN 21726 21766  
FT DOMAIN 21766 21806  
FT DOMAIN 21806 21846  
FT DOMAIN 21846 21886  
FT DOMAIN 21886 21926  
FT DOMAIN 21926 21966  
FT DOMAIN 21966 22006  
FT DOMAIN 22006 22046  
FT DOMAIN 22046 22086  
FT DOMAIN 22086 22126  
FT DOMAIN 22126 22166  
FT DOMAIN 22166 22206  
FT DOMAIN 22206 22246  
FT DOMAIN 22246 22286  
FT DOMAIN 22286 22326  
FT DOMAIN 22326 22366  
FT DOMAIN 22366 22406  
FT DOMAIN 22406 22446  
FT DOMAIN 22446 22486  
FT DOMAIN 22486 22526  
FT DOMAIN 22526 22566  
FT DOMAIN 22566 22606  
FT DOMAIN 22606 22646  
FT DOMAIN 22646 22686  
FT DOMAIN 22686 22726  
FT DOMAIN 22726 22766  
FT DOMAIN 22766 22806  
FT DOMAIN 22806 22846  
FT DOMAIN 22846 22886  
FT DOMAIN 22886 22926  
FT DOMAIN 22926 22966  
FT DOMAIN 22966 23006  
FT DOMAIN 23006 23046  
FT DOMAIN 23046 23086  
FT DOMAIN 23086 23126  
FT DOMAIN 23126 23166  
FT DOMAIN 23166 23206  
FT DOMAIN 23206 23246  
FT DOMAIN 23246 23286  
FT DOMAIN 23286 23326  
FT DOMAIN 23326 23366  
FT DOMAIN 23366 23406  
FT DOMAIN 23406 23446  
FT DOMAIN 23446 23486  
FT DOMAIN 23486 23526  
FT DOMAIN 23526 23566  
FT DOMAIN 23566 23606  
FT DOMAIN 23606 23646  
FT DOMAIN 23646 23686  
FT DOMAIN 23686 23726  
FT DOMAIN 23726 23766  
FT DOMAIN 23766 23806  
FT DOMAIN 23806 23846  
FT DOMAIN 23846 23886  
FT DOMAIN 23886 23926  
FT DOMAIN 23926 23966  
FT DOMAIN 23966 24006  
FT DOMAIN 24006 24046  
FT DOMAIN 24046 24086  
FT DOMAIN 24086 24126  
FT DOMAIN 24126 24166  
FT DOMAIN 24166 24206  
FT DOMAIN 24206 24246  
FT DOMAIN 24246 24286  
FT DOMAIN 24286 24326  
FT DOMAIN 24326 24366  
FT DOMAIN 24366 24406  
FT DOMAIN 24406 24446  
FT DOMAIN 24446 24486  
FT DOMAIN 24486 24526  
FT DOMAIN 24526 24566  
FT DOMAIN 24566 24606  
FT DOMAIN 24606 24646  
FT DOMAIN 24646 24686  
FT DOMAIN 24686 24726  
FT DOMAIN 24726 24766  
FT DOMAIN 24766 24806  
FT DOMAIN 24806 24846  
FT DOMAIN 24846 24886  
FT DOMAIN 24886 24926  
FT DOMAIN 24926 24966  
FT DOMAIN 24966 25006  
FT DOMAIN 25006 25046  
FT DOMAIN 25046 25086  
FT DOMAIN 25086 25126  
FT DOMAIN 25126 25166  
FT DOMAIN 25166 25206  
FT DOMAIN 25206 25246  
FT DOMAIN 25246 25286  
FT DOMAIN 25286 25326  
FT DOMAIN 25326 25366  
FT DOMAIN 25366 25406  
FT DOMAIN 25406 25446  
FT DOMAIN 25446 25486  
FT DOMAIN 25486 25526  
FT DOMAIN 25526 25566  
FT DOMAIN 25566 25606  
FT DOMAIN 25606 25646  
FT DOMAIN 25646 25686  
FT DOMAIN 25686 25726  
FT DOMAIN 25726 25766  
FT DOMAIN 25766 25806  
FT DOMAIN 25806 25846  
FT DOMAIN 25846 25886  
FT DOMAIN 25886 25926  
FT DOMAIN 25926 25966  
FT DOMAIN 25966 26006  
FT DOMAIN 26006 26046  
FT DOMAIN 26046 26086  
FT DOMAIN 26086 26126  
FT DOMAIN 26126 26166  
FT DOMAIN 26166 26206  
FT DOMAIN 26206 26246  
FT DOMAIN 26246 26286  
FT DOMAIN 26286 26326  
FT DOMAIN 26326 26366  
FT DOMAIN 26366 26406  
FT DOMAIN 26406 26446  
FT DOMAIN 26446 26486  
FT DOMAIN 26486 26526  
FT DOMAIN 26526 26566  
FT DOMAIN 26566 26606  
FT DOMAIN 26606 26646  
FT DOMAIN 26646 26686  
FT DOMAIN 26686 26726  
FT DOMAIN 26726 26766  
FT DOMAIN 26766 26806  
FT DOMAIN 26806 26846  
FT DOMAIN 26846 26886  
FT DOMAIN 26886 26926  
FT DOMAIN 26926 26966  
FT DOMAIN 26966 27006  
FT DOMAIN 27006 27046  
FT DOMAIN 27046 27086  
FT DOMAIN 27086 27126  
FT DOMAIN 27126 27166  
FT DOMAIN 27166 27206  
FT DOMAIN 27206 27246  
FT DOMAIN 27246 27286  
FT DOMAIN 27286 27326  
FT DOMAIN 27326 27366  
FT DOMAIN 27366 27406  
FT DOMAIN 27406 27446  
FT DOMAIN 27446 27486  
FT DOMAIN 27486 27526  
FT DOMAIN 27526 27566  
FT DOMAIN 27566 27606  
FT DOMAIN 27606 27646  
FT DOMAIN 27646 27686  
FT DOMAIN 27686 27726  
FT DOMAIN 27726 277



```
FT STRAND 46 48
FT TURN 50 51
FT STRAND 54 56
FT HELIX 57 59
SQ SEQUENCE 63 AA; 6862 MW; 5DC10DE75B375F16 CRC64;

Query Match 19.0%; Score 92.5; DB 1; Length 63;
Best Local Similarity 32.8%; Pred. No. 0.0076;
Matches 22; Conservative 5; Mismatches 27; Indels 13; Gaps 4;

QY 6 CGENKDYDCGSKCKKCYDCVEEDDEPNVCLVRVCHQDCVCE--EGFYRNKDK 63
Db 5 CGNEVWTECTG--CEMKC-----NTPCLMCRPSCSPGRCGRRTNDGK 53

QY 64 CVAEDC 70
Db 54 CIPASQC 60

RESULT 7
LRP CAEL STANDARD; PRT; 4753 AA.
AC Q04833;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein precursor (LRP).
GN LRP-1 OR F29D11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281621; PubMed=8506301;
RA Yochem J.; Greenwald I.;
RT "A gene for a low density lipoprotein receptor-related protein in the
RT nematode Caenorhabditis elegans.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:4572-4576(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilkinson J.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY ACT AS A RECEPTOR FOR THE ENDOCYTOSIS OF
CC EXTRACELLULAR LIGANDS SUCH AS CHYLOMICRON REMNANTS, PROTEASE-
CC INHIBITOR COMPLEXES AND VITELLOGENIN.
CC -!- SUBCELLULAR LOCATION: type I membrane protein.
CC -!- SIMILARITY: Contains 35 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 17 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M96150; AAA28105.1; -
CC DR EMBL; 273907; CAA98124.1; -
CC DR PIR; A47437; A47437.
CC DR HSP; Q07954; 1CR8.
CC DR WormPep; F29D11.1; CE05765.
CC DR InterPro; IPR000152; Asx_hydroxyl.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR002172; LDL_receptor_A.
CC DR InterPro; IPR000033; Ldl_receptor_rep.
CC DR Pfam; PF00057; ldl_recept_a; 34.
CC DR Pfam; PF00058; ldl_recept_b; 26.
CC DR PRINTS; PR00261; LDLRECEPTOR.
CC DR SMART; SM00179; EGF_CA; 2.
```

```
DR SMART; SM00192; LDLa; 35.
DR SMART; SM00135; LY; 32.
DR PROSITE; PS00010; ASX_HYDROXYL; 6.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 34.
KW Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein;
KW Signal; Calcium-binding; EGF-like domain; Coated pits.
FT SIGNAL 1 18
FT CHAIN 19 4753
FT DOMAIN 19 4570
FT TRANSMEM 4571 4596
FT DOMAIN 4597 4753
FT DOMAIN 51 89
FT DOMAIN 90 133
FT DOMAIN 136 177
FT DOMAIN 180 220
FT DOMAIN 221 259
FT DOMAIN 260 298
FT DOMAIN 299 337
FT DOMAIN 338 368
FT DOMAIN 669 712
FT DOMAIN 997 1043
FT DOMAIN 1052 1097
FT DOMAIN 1144 1184
FT DOMAIN 1185 1225
FT DOMAIN 1226 1265
FT DOMAIN 1268 1309
FT DOMAIN 1311 1352
FT DOMAIN 1357 1397
FT DOMAIN 1398 1436
FT DOMAIN 1437 1476
FT DOMAIN 1747 1786
FT DOMAIN 2080 2120
FT DOMAIN 2396 2439
FT DOMAIN 2728 2780
FT DOMAIN 2790 2831
FT DOMAIN 2832 2870
FT DOMAIN 2872 2914
FT DOMAIN 2917 2958
FT DOMAIN 2959 2999
FT DOMAIN 3004 3046
FT DOMAIN 3047 3095
FT DOMAIN 3137 3176
FT DOMAIN 3185 3223
FT DOMAIN 3224 3265
FT DOMAIN 3266 3306
FT DOMAIN 3382 3424
FT DOMAIN 3625 3668
FT DOMAIN 3669 3707
FT DOMAIN 3707 3748
FT DOMAIN 3751 3790
FT DOMAIN 3791 3832
FT DOMAIN 3831 3873
FT DOMAIN 3876 3914
FT DOMAIN 3915 3953
FT DOMAIN 3957 3997
FT DOMAIN 3998 4042
FT DOMAIN 4047 4085
FT DOMAIN 4088 4131
FT DOMAIN 4132 4176
FT DOMAIN 4477 4515
FT DOMAIN 4526 4554
FT SITE 4553 4558
FT SITE 4744 4744
FT SITE 53 65
FT DISULFID 60 78
FT DISULFID 72 87
```



```

CC DR EMBL; L76227; AAB05549.1; -
DR DR EMBL; L16903; AAA30903.1; -
DR DR EMBL; AF099154; AAD04919.1; -
DR DR EMBL; U66246; AAB93766.1; -
DR DR HSSP; P04275; 1AUQ.
DR DR InterPro; IPR006208; Cys_knot.
DR DR InterPro; IPR006207; Cys_knot_C.
DR DR InterPro; IPR000004; SapB.
DR DR InterPro; IPR002919; TIL_Cysrich.
DR DR InterPro; IPR002035; VWF_A.
DR DR InterPro; IPR001007; VWF_C.
DR DR InterPro; IPR001846; VWF_D.
DR DR Pfam; PF00007; Cys_knot; 1.
DR DR Pfam; PF01826; TIL; 4.
DR DR Pfam; PF00092; vwa; 3.
DR DR Pfam; PF00093; vwc; 3.
DR DR Pfam; PF00094; vwg; 4.
DR DR PRINTS; SM00453; VWFADOMAIN.
DR DR SMART; SM00041; CT; 1.
DR DR SMART; SM00327; VWA; 3.
DR DR SMART; SM00214; VWC; 5.
DR DR SMART; SM00216; VWD; 4.
DR DR PROSITE; PS01185; CTCK_1; 1.
DR DR PROSITE; PS01225; CTCK_2; 1.
DR DR PROSITE; PS0234; VWFA; 3.
DR DR PROSITE; PS01208; VWFC_1; 3.
DR DR PROSITE; PS0184; VWFC_2; 3.
KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Plasma; Hemostasis; Repeat; Cell adhesion; Signal.
FT SIGNAL 1 22
FT PROPEP 23 763
FT CHAIN 764 2813
FT DOMAIN 35 179
FT DOMAIN 388 541
FT DOMAIN 764 787
FT DOMAIN 788 833
FT DOMAIN 826 853
FT DOMAIN 867 1013
FT DOMAIN 1277 1453
FT DOMAIN 1498 1665
FT DOMAIN 1691 1871
FT DOMAIN 1950 2102
FT DOMAIN 2216 2261
FT DOMAIN 2255 2328
FT DOMAIN 2429 2495
FT DOMAIN 2580 2645
FT DOMAIN 2724 2812
FT SITE 531 533
FT SITE 698 700
FT SITE 2507 2509
FT DISULFID 767 808
FT DISULFID 776 804
FT DISULFID 810 821
FT DISULFID 867 996
FT DISULFID 889 1031
FT DISULFID 898 993
FT DISULFID 914 921
FT DISULFID 1060 1084
FT DISULFID 1071 1111
FT DISULFID 1089 1091
FT DISULFID 1153 1165
FT DISULFID 1149 1169
FT DISULFID 1126 1130
FT DISULFID 1136 1199
FT DISULFID 1234 1237
FT DISULFID 1272 1458
FT DISULFID 1669 1670
FT DISULFID 1686 1872
FT DISULFID 1879 1904
FT DISULFID 1899 1940
FT DISULFID 1972 2123
FT DISULFID 1950 2085
DR EMBL; L76227; AAB05549.1; -
DR EMBL; L16903; AAA30903.1; -
DR EMBL; AF099154; AAD04919.1; -
DR EMBL; U66246; AAB93766.1; -
DR HSSP; P04275; 1AUQ.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000004; SapB.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR002035; VWF_A.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF01826; TIL; 4.
DR Pfam; PF00092; vwa; 3.
DR Pfam; PF00093; vwc; 3.
DR Pfam; PF00094; vwg; 4.
DR PRINTS; SM00453; VWFADOMAIN.
DR SMART; SM00041; CT; 1.
DR SMART; SM00327; VWA; 3.
DR SMART; SM00214; VWC; 5.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS0234; VWFA; 3.
DR PROSITE; PS01208; VWFC_1; 3.
DR PROSITE; PS0184; VWFC_2; 3.
KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Plasma; Hemostasis; Repeat; Cell adhesion; Signal.
FT SIGNAL 1 22
FT PROPEP 23 763
FT CHAIN 764 2813
FT DOMAIN 35 179
FT DOMAIN 388 541
FT DOMAIN 764 787
FT DOMAIN 788 833
FT DOMAIN 826 853
FT DOMAIN 867 1013
FT DOMAIN 1277 1453
FT DOMAIN 1498 1665
FT DOMAIN 1691 1871
FT DOMAIN 1950 2102
FT DOMAIN 2216 2261
FT DOMAIN 2255 2328
FT DOMAIN 2429 2495
FT DOMAIN 2580 2645
FT DOMAIN 2724 2812
FT SITE 531 533
FT SITE 698 700
FT SITE 2507 2509
FT DISULFID 767 808
FT DISULFID 776 804
FT DISULFID 810 821
FT DISULFID 867 996
FT DISULFID 889 1031
FT DISULFID 898 993
FT DISULFID 914 921
FT DISULFID 1060 1084
FT DISULFID 1071 1111
FT DISULFID 1089 1091
FT DISULFID 1153 1165
FT DISULFID 1149 1169
FT DISULFID 1126 1130
FT DISULFID 1136 1199
FT DISULFID 1234 1237
FT DISULFID 1272 1458
FT DISULFID 1669 1670
FT DISULFID 1686 1872
FT DISULFID 1879 1904
FT DISULFID 1899 1940
FT DISULFID 1972 2123
FT DISULFID 1950 2085

```

```

FT DISULFID 1927 2088
FT DISULFID 1993 2001
FT DISULFID 2724 2774
FT DISULFID 2739 2788
FT DISULFID 2750 2804
FT DISULFID 2754 2806
FT DISULFID ? 2811
FT CARBOHYD 99 99
FT CARBOHYD 156 156
FT CARBOHYD 211 211
FT CARBOHYD 666 666
FT CARBOHYD 857 857
FT CARBOHYD 1231 1231
FT CARBOHYD 1515 1515
FT CARBOHYD 1574 1574
FT CARBOHYD 2223 2223
FT CARBOHYD 2290 2290
FT CARBOHYD 2357 2357
FT CARBOHYD 2400 2400
FT CARBOHYD 2546 2546
FT CARBOHYD 2585 2585
FT CARBOHYD 2790 2790
FT CONFLICT 55 55
FT CONFLICT 70 70
FT CONFLICT 266 266
FT CONFLICT 280 280
FT CONFLICT 409 411
FT CONFLICT 994 994
FT CONFLICT 1021 1021
FT CONFLICT 2381 2381
FT CONFLICT 2406 2406
SQ SEQUENCE 2813 AA; 309716 MW; 5DF93E1E5E72F80C CRC64;

Query Match 18.0%; Score 87.5; DB 1; Length 2813;
Best Local Similarity 28.4%; Pred. No. 1;
Matches 19; Conservative 9; Mismatches 30; Indels 9; Gaps 2;

QY 4 MQGGENEKYDSCGKCEKCKYDGVBEDEEPNVLVRVCHDVCCEGFYRNKDDK 63
DB 650 LSCPGQVYLOQGT-PCNWTCSRSLSYPEDCNE-----VCLGGCFPPGLYDGRGD 700

QY 64 CVSAEDC 70
DB 701 CVPKAQ 707

RESULT 9
AG2F DROME
ID AG2F DROME STANDARD; PRT; 115 AA.
AC 046202;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Accessory gland protein Acp62F precursor.
GN ACP62F OR CGI262.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Male accessory gland;
RX MEDLINE=98135120; PubMed=9474779;
RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,
RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;
RT "New genes for male accessory gland proteins in Drosophila
RT melanogaster."
RL Insect Biochem. Mol. Biol. 27:825-834(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;

```

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scher S.E., Li P.W., Hoekins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Buzam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foaier C., Gabriellian A.C., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2185-2195 (2000).  
[3]  
RP SEQUENCE OF 7-111 FROM N.A.  
RC STRAIN=ZIM62H-12C, ZIM62H-16C, ZIM62H-28C, ZIM62H-30C, ZIM62H-34C,  
RC ZIM62I-5C, ZIM62I-10C, ZIM62I-17C, ZIM62I-18C, and ZIM62I-53C;  
RX MEDLINE=20556153; PubMed=11102381;  
RA Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;  
RT "Molecular population genetics of male accessory gland proteins in  
RT *Drosophila*";  
RL Genetics 156:1879-1888 (2000).  
CC -!- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
CC MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL  
CC FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO  
CC AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE  
CC AND EGG RELEASE.  
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -!- TISSUE SPECIFICITY: SEMINAL FLUID.  
CC -!- SIMILARITY: SOME, TO P.NIGRIVENTER TX2-6.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; U85763; AAB96387.1; -;  
DR EMBL; AE003475; AAF47683.1; -;  
DR EMBL; AY010608; AAG35367.1; -;  
DR EMBL; AY010609; AAG35368.1; -;  
DR EMBL; AY010610; AAG35369.1; -;  
DR EMBL; AY010611; AAG35370.1; -;  
DR EMBL; AY010612; AAG35371.1; -;  
DR EMBL; AY010613; AAG35372.1; -;  
DR EMBL; AY010614; AAG35373.1; -;

DR EMBL; AY010615; AAG35374.1; -;  
DR EMBL; AY010616; AAG35375.1; -;  
DR EMBL; AY010617; AAG35376.1; -;  
DR FlyBase; FBgn0020509; Acp62F.  
DR GO; GO:0004867; F:serine protease inhibitor activity; IDA.  
DR InterPro; IPR002919; TIL\_Cysrich.  
DR Pfam; PF01826; TIL; 1.  
KW Behavior; Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 115 ACCESSORY GLAND PROTEIN ACP62F.  
FT DOMAIN 34 88 TIL.  
SQ SEQUENCE 115 AA; 12570 MW; 4326AA6F6C32291D CRC64;  
  
Query Match 17.9%; Score 87; DB 1; Length 115;  
Best Local Similarity 31.0%; Pred. No. 0.047;  
Matches 22; Conservative 8; Mismatches 29; Indels 12; Gaps 4;  
  
QY 4 MCGENKDYDSGSKCKKCKYGVVEEDDEPNVPCLVRVCHQDCVCEGFPYRN-KDD 62  
Db 32 VDCNTANGTQTEC-PVACPETCEYSG-----NGPC-VKMGAPCVCKPGVIVNERIP 80  
QY 63 KCVSAEDCSELD 73  
Db 81 ACVLRSDCPKD 91  
  
RESULT 10  
LRP2\_HUMAN STANDARD; PRT; 4655 AA.  
ID AC P98164; O00711; Q16215;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)  
DE (Glycoprotein 330) (gp330).  
GN LRP2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=963105376; PubMed=8706697;  
RA Hjaelm G., Murray E., Crumley G., Harazim W., Lundgren S., Onyango I.,  
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstrom G.,  
RA Rask L., Morse B.;  
RT "Cloning and sequencing of human gp330, a Ca(2+)-binding receptor  
RT with potential intracellular signaling properties";  
RL Eur. J. Biochem. 239:132-137 (1996).  
RN [2]  
RP SEQUENCE OF 2705-4453 FROM N.A.  
RC TISSUE=Kidney;  
RA Knaak C., Argraves W.S.;  
RN Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 3833-4453 FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=95048397; PubMed=7959795;  
RA Korenberg J.R., Argraves K.M., Chen X.N., Tran H.,  
RA Strickland D.K., Argraves W.S.;  
RT "Chromosomal localization of human genes for the LDL receptor family  
RT member glycoprotein 330 (LRP2) and its associated protein RAP  
RT (LRPAP1).";  
RL Genomics 22:88-93 (1994).  
RN [4]  
RP SEQUENCE OF 4139-4406 FROM N.A.  
RX MEDLINE=94244704; PubMed=8187828;  
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,  
RA Klarskog L., Akerstrom G., Rask L.;  
RT "A protein involved in calcium sensing of the human parathyroid and  
RT placental cytotrophoblast cells belongs to the LDL-receptor protein  
RT superfamily";

Exp. Cell Res. 212:344-350(1994).  
[5]  
FUNCTION.  
RA MEDLINE=95286588; PubMed=7768901;  
RX Kounnas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,  
RA Brewer B.H., Strickland D.K., Agraves W.S.;  
RT Identification of glycoprotein 330 as an endocytic receptor for  
apoptosis-inducing factor.  
J. Biol. Chem. 270:13070-13075(1995).  
CC -!- FUNCTION: BINDS SPECIFICALLY CLUSTERIN WITH HIGH AFFINITY. BUT  
CC ALSO LIGANDS IN COMMON WITH OTHER FAMILY MEMBERS: PLASMINOGEN  
CC EXTRACELLULAR MATRIX COMPONENTS, PLASMINOGEN ACTIVATOR-PLASMINOGEN  
CC ACTIVATOR INHIBITOR TYPE 1 COMPLEX, APOLOPROTEIN E-ENRICHED  
CC BETA-VLDL, LIPOPROTEIN LIPASE, LACTOFERRIN AND CALCIUM.  
CC -!- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS  
CC APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B (BY SIMILARITY).  
CC -!- FUNCTION: MAY PARTICIPATE IN REGULATION OF PARATHYROID-HORMONE AND  
CC PARA-THYROID-HORMONE-RELATED PROTEIN RELEASE.  
CC -!- SUBUNIT: Forms a multimeric complex together with a receptor-  
CC associated protein (RAP). Binds to ankyrin-repeat family A protein  
CC 2 (ANKRA2) (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: ABSORPTIVE EPITHELIA, INCLUDING RENAL  
CC PROXIMAL TUBULES.  
CC -!- SIMILARITY: Contains 36 LDL-receptor class A domains.  
CC -!- SIMILARITY: Contains 37 LDL-receptor class B domains.  
CC -!- SIMILARITY: Contains 17 EGF-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; U33837; AAB41649.1; -;  
DR EMBL; U04441; AAB02882.1; -;  
DR EMBL; S73145; AAB30825.1; -;  
DR PIR; I53413; I53413.  
DR HSRP; Q07954; ICR8.  
DR Genew; HGNC:6694; LRP2.  
DR MIM; 600073; -;  
DR GO; GO:0005764; C:lysosome; TAS.  
DR GO; GO:0006629; P:lipid metabolism; TAS.  
DR GO; GO:0006486; P:protein amino acid glycosylation; TAS.  
DR GO; GO:0006898; P:receptor mediated endocytosis; TAS.  
DR InterPro; IPR000152; Abx\_hydroxyl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; LDL\_receptor\_rep.  
DR Pfam; PF00008; EGF; 13.  
DR Pfam; PF00057; ldl\_recept\_a; 36.  
DR Pfam; PF00058; ldl\_recept\_b; 37.  
DR SMART; SM00179; EGF\_CA; 3.  
DR SMART; SM00192; LDLA; 36.  
DR SMART; SM00135; LV; 33.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 9.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLRA\_1; 31.  
DR PROSITE; PS00688; LDLRA\_2; 36.  
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;  
KW Receptor; EGF-like domain; Signal; SH3-binding; Polymorphism.  
FT SIGNAL 1 25  
FT CHAIN 26 4655 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED  
FT PROTEIN 2.  
FT DOMAIN 26 4423 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 4424 4446 POTENTIAL.  
FT DOMAIN 4447 4655 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 26 64 LDL-RECEPTOR CLASS A 1.

65 105  
106 144  
145 181  
182 219  
220 258  
259 308  
309 347  
348 386  
387 425  
426 464  
465 503  
504 542  
543 581  
582 620  
621 659  
660 700  
701 740  
741 780  
781 820  
821 860  
861 900  
901 940  
941 980  
981 1020  
1021 1060  
1061 1100  
1101 1140  
1141 1180  
1181 1220  
1221 1260  
1261 1300  
1301 1340  
1341 1380  
1381 1420  
1421 1460  
1461 1500  
1501 1540  
1541 1580  
1581 1620  
1621 1660  
1661 1700  
1701 1740  
1741 1780  
1781 1820  
1821 1860  
1861 1900  
1901 1940  
1941 1980  
1981 2020  
2021 2060  
2061 2100  
2101 2140  
2141 2180  
2181 2220  
2221 2260  
2261 2300  
2301 2340  
2341 2380  
2381 2420  
2421 2460  
2461 2500  
2501 2540  
2541 2580  
2581 2620  
2621 2660  
2661 2700  
2701 2740  
2741 2780  
2781 2820  
2821 2860  
2861 2900  
2901 2940  
2941 2980  
2981 3020  
3021 3060  
3061 3100  
3101 3140  
3141 3180  
3181 3220  
3221 3260  
3261 3300  
3301 3340  
3341 3380  
3381 3420  
3421 3460  
3461 3500  
3501 3540  
3541 3580  
3581 3620  
3621 3660  
3661 3700  
3701 3740  
3741 3780  
3781 3820  
3821 3860  
3861 3900  
3901 3940  
3941 3980  
3981 4020  
4021 4060  
4061 4100  
4101 4140  
4141 4180  
4181 4220  
4221 4260  
4261 4300  
4301 4340  
4341 4380  
4381 4420  
4421 4460  
4461 4500  
4501 4540  
4541 4580  
4581 4620  
4621 4660  
4661 4700  
4701 4740  
4741 4780  
4781 4820  
4821 4860  
4861 4900  
4901 4940  
4941 4980  
4981 5020  
5021 5060  
5061 5100  
5101 5140  
5141 5180  
5181 5220  
5221 5260  
5261 5300  
5301 5340  
5341 5380  
5381 5420  
5421 5460  
5461 5500  
5501 5540  
5541 5580  
5581 5620  
5621 5660  
5661 5700  
5701 5740  
5741 5780  
5781 5820  
5821 5860  
5861 5900  
5901 5940  
5941 5980  
5981 6020  
6021 6060  
6061 6100  
6101 6140  
6141 6180  
6181 6220  
6221 6260  
6261 6300  
6301 6340  
6341 6380  
6381 6420  
6421 6460  
6461 6500  
6501 6540  
6541 6580  
6581 6620  
6621 6660  
6661 6700  
6701 6740  
6741 6780  
6781 6820  
6821 6860  
6861 6900  
6901 6940  
6941 6980  
6981 7020  
7021 7060  
7061 7100  
7101 7140  
7141 7180  
7181 7220  
7221 7260  
7261 7300  
7301 7340  
7341 7380  
7381 7420  
7421 7460  
7461 7500  
7501 7540  
7541 7580  
7581 7620  
7621 7660  
7661 7700  
7701 7740  
7741 7780  
7781 7820  
7821 7860  
7861 7900  
7901 7940  
7941 7980  
7981 8020  
8021 8060  
8061 8100  
8101 8140  
8141 8180  
8181 8220  
8221 8260  
8261 8300  
8301 8340  
8341 8380  
8381 8420  
8421 8460  
8461 8500  
8501 8540  
8541 8580  
8581 8620  
8621 8660  
8661 8700  
8701 8740  
8741 8780  
8781 8820  
8821 8860  
8861 8900  
8901 8940  
8941 8980  
8981 9020  
9021 9060  
9061 9100  
9101 9140  
9141 9180  
9181 9220  
9221 9260  
9261 9300  
9301 9340  
9341 9380  
9381 9420  
9421 9460  
9461 9500  
9501 9540  
9541 9580  
9581 9620  
9621 9660  
9661 9700  
9701 9740  
9741 9780  
9781 9820  
9821 9860  
9861 9900  
9901 9940  
9941 9980  
9981 10020  
10021 10060  
10061 10100  
10101 10140  
10141 10180  
10181 10220  
10221 10260  
10261 10300  
10301 10340  
10341 10380  
10381 10420  
10421 10460  
10461 10500  
10501 10540  
10541 10580  
10581 10620  
10621 10660  
10661 10700  
10701 10740  
10741 10780  
10781 10820  
10821 10860  
10861 10900  
10901 10940  
10941 10980  
10981 11020  
11021 11060  
11061 11100  
11101 11140  
11141 11180  
11181 11220  
11221 11260  
11261 11300  
11301 11340  
11341 11380  
11381 11420  
11421 11460  
11461 11500  
11501 11540  
11541 11580  
11581 11620  
11621 11660  
11661 11700  
11701 11740  
11741 11780  
11781 11820  
11821 11860  
11861 11900  
11901 11940  
11941 11980  
11981 12020  
12021 12060  
12061 12100  
12101 12140  
12141 12180  
12181 12220  
12221 12260  
12261 12300  
12301 12340  
12341 12380  
12381 12420  
12421 12460  
12461 12500  
12501 12540  
12541 12580  
12581 12620  
12621 12660  
12661 12700  
12701 12740  
12741 12780  
12781 12820  
12821 12860  
12861 12900  
12901 12940  
12941 12980  
12981 13020  
13021 13060  
13061 13100  
13101 13140  
13141 13180  
13181 13220  
13221 13260  
13261 13300  
13301 13340  
13341 13380  
13381 13420  
13421 13460  
13461 13500  
13501 13540  
13541 13580  
13581 13620  
13621 13660  
13661 13700  
13701 13740  
13741 13780  
13781 13820  
13821 13860  
13861 13900  
13901 13940  
13941 13980  
13981 14020  
14021 14060  
14061 14100  
14101 14140  
14141 14180  
14181 14220  
14221 14260  
14261 14300  
14301 14340  
14341 14380  
14381 14420  
14421 14460  
14461 14500  
14501 14540  
14541 14580  
14581 14620  
14621 14660  
14661 14700  
14701 14740  
14741 14780  
14781 14820  
14821 14860  
14861 14900  
14901 14940  
14941 14980  
14981 15020  
15021 15060  
15061 15100  
15101 15140  
15141 15180  
15181 15220  
15221 15260  
15261 15300  
15301 15340  
15341 15380  
15381 15420  
15421 15460  
15461 15500  
15501 15540  
15541 15580  
15581 15620  
15621 15660  
15661 15700  
15701 15740  
15741 15780  
15781 15820  
15821 15860  
15861 15900  
15901 15940  
15941 15980  
15981 16020  
16021 16060  
16061 16100  
16101 16140  
16141 16180  
16181 16220  
16221 16260  
16261 16300  
16301 16340  
16341 16380  
16381 16420  
16421 16460  
16461 16500  
16501 16540  
16541 16580  
16581 16620  
16621 16660  
16661 16700  
16701 16740  
16741 16780  
16781 16820  
16821 16860  
16861 16900  
16901 16940  
16941 16980  
16981 17020  
17021 17060  
17061 17100  
17101 17140  
17141 17180  
17181 17220  
17221 17260  
17261 17300  
17301 17340  
17341 17380  
17381 17420  
17421 17460  
17461 17500  
17501 17540  
17541 17580  
17581 17620  
17621 17660  
17661 17700  
17701 17740  
17741 17780  
17781 17820  
17821 17860  
17861 17900  
17901 17940  
17941 17980  
17981 18020  
18021 18060  
18061 18100  
18101 18140  
18141 18180  
18181 18220  
18221 18260  
18261 18300  
18301 18340  
18341 18380  
18381 18420  
18421 18460  
18461 18500  
18501 18540  
18541 18580  
18581 18620  
18621 18660  
18661 18700  
18701 18740  
18741 18780  
18781 18820  
18821 18860  
18861 18900  
18901 18940  
18941 18980  
18981 19020  
19021 19060  
19061 19100  
19101 19140  
19141 19180  
19181 19220  
19221 19260  
19261 19300  
19301 19340  
19341 19380  
19381 19420  
19421 19460  
19461 19500  
19501 19540  
19541 19580  
19581 19620  
19621 19660  
19661 19700  
19701 19740  
19741 19780  
19781 19820  
19821 19860  
19861 19900  
19901 19940  
19941 19980  
19981 20020  
20021 20060  
20061 20100  
20101 20140  
20141 20180  
20181 20220  
20221 20260  
20261 20300  
20301 20340  
20341 20380  
20381 20420  
20421 20460  
20461 20500  
20501 20540  
20541 20580  
20581 20620  
20621 20660  
20661 20700  
20701 20740  
20741 20780  
20781 20820  
20821 20860  
20861 20900  
20901 20940  
20941 20980  
20981 21020  
21021 21060  
21061 21100  
21101 21140  
21141 21180  
21181 21220  
21221 21260  
21261 21300  
21301 21340  
21341 21380  
21381 21420  
21421 21460  
21461 21500  
21501 21540  
21541 21580  
21581 21620  
21621 21660  
21661 21700  
21701 21740  
21741 21780  
21781 21820  
21821 21860  
21861 21900  
21901 21940  
21941 21980  
21981 22020  
22021 22060  
22061 22100  
22101 22140  
22141 22180  
22181 22220  
22221 22260  
22261 22300  
22301 22340  
22341 22380  
22381 22420  
22421 22460  
22461 22500  
22501 22540  
22541 22580  
22581 22620  
22621 22660  
22661 22700  
22701 22740  
22741 22780  
22781 22820  
22821 22860  
22861 22900  
22901 22940  
22941 22980  
22981 23020  
23021 23060  
23061 23100  
23101 23140  
23141 23180  
23181 23220  
23221 23260  
23261 23300  
23301 23340  
23341 23380  
23381 23420  
23421 23460  
23461 23500  
23501 23540  
23541 23580  
23581 23620  
23621 23660  
23661 23700  
23701 23740  
23741 23780  
23781 23820  
23821 23860  
23861 23900  
23901 23940  
23941 23980  
23981 24020  
24021 24060  
24061 24100  
24101 24140  
24141 24180  
24181 24220  
24221 24260  
24261 24300  
24301 24340  
24341 24380  
24381 24420  
24421 24460  
24461 24500  
24501 24540  
24541 24580  
24581 24620  
24621 24660  
24661 24700  
24701 24740  
24741 24780  
24781 24820  
24821 24860  
24861 24900  
24901 24940  
24941 24980  
24981 25020  
25021 25060  
25061 25100  
25101 25140  
25141 25180  
25181 25220  
25221 25260  
25261 25300  
25301 25340  
25341 25380  
25381 25420  
25421 25460  
25461 25500  
25501 25540  
25541 25580  
25581 25620  
25621 25660  
25661 25700  
25701 25740  
25741 25780  
25781 25820  
25821 25860  
25861 25900  
25901 25940  
25941 25980  
25981 26020  
26021 26060  
26061 26100  
26101 26140  
26141 26180  
26181 26220  
26221 26260  
26261 26300  
26301 26340  
26341 26380  
26381 26420  
26421 26460  
26461 26500  
26501 26540  
26541 26580  
26581 26620  
26621 26660  
26661 26700  
26701 26740  
26741 26780  
26781 26820  
26821 26860  
26861 26900  
26901 26940  
26941 26980  
26981 27020  
27021 27060  
27061 27100  
27101 27140  
27141 27180  
27181 27220  
27221 27260  
27261 27300  
27301 27340  
27341 27380  
27381 27420  
27421 27460  
27461 27500  
27501 27540  
27541 27580  
27581 27620  
27621 27660  
27661 27700  
27701 27740  
27741 27780  
27781 27820  
27821 27860  
27861 27900  
27901 27940  
27941 27980  
27981 28020  
28021 28060  
28061 28100  
28101 28140  
28141 28180  
28181 28220  
28221 28260  
28261 28300  
28301 28340  
28341 28380  
28381 28420  
28421 28460  
28461 28500  
28501 28540  
28541 28580  
28581 28620  
28621 28660  
28661 28700  
28701 28740  
28741 28780  
28781 28820  
28821 28860  
28861 28900  
28901 28940  
28941 28980  
28981 29020  
29021 29060  
29061 29100  
29101 29140  
29141 29180  
29181 29220  
29221 29260  
29261 29300  
29301 29340  
29341 29380  
29381 29420  
29421 29460  
29461 29500  
29501 29540  
29541 29580  
29581 29620  
29621 29660  
29661 29700  
29701 29740  
29741 29780  
29781 29820  
29821 29860  
29861 29900  
29901 29940  
29941 29980  
29981 30020  
30021 30060  
30061 30100  
30101 30140  
30141 30180  
30181 30220  
30221 30260  
30261 30300  
30301 30340  
30341 30380  
30381 30420  
30421 30460  
30461 30500  
30501 30540  
30541 30580  
30581 30620  
30621 30660  
30661 30700  
30701 30740  
30741 30780  
30781 30820  
30821 30860  
30861 30900  
30901 30940  
30941 30980  
30981 31020  
31021 31060  
31061 31100  
31101 31140  
31141 31180  
31181 31220  
31221 31260  
31261 31300  
31301 31340  
31341 31380  
31381 31420  
31421 31460  
31461 31500  
31501 31540  
31541 31580  
31581 31620  
31621 31660  
31661 31700  
31701 31740  
31741 31780  
31781 31820  
31821 31860  
31861 31900  
31901 31940  
31941 31980  
31981 32020  
32021 32060  
32061 32100  
32101 32140  
32141 32180  
32181 32220  
32221 32260  
32261 32300  
32301 32340  
32341 32380  
32381 32420  
32421 32460  
32461 32500  
32501 32540  
32541 32580  
32581 32620  
32621 32660  
32661 32700  
32701 32740  
32741 32780  
32781 32820  
32821 32860  
32861 32900  
32901 32940  
32941 32980  
32981 33020  
33021 33060  
33061 33100  
33101 33140  
33141 33180  
33181 33220  
33221 33260  
33261 33300  
33301 33340  
33341 33380  
33381 33420  
33421 33460  
33461 33500  
33501 33540  
33541 33580  
33581 33620  
33621 33660  
33661 33700  
33701 33740  
33741 33780  
33781 33820  
33821 33860  
33861 33900  
33901 33940  
33941 33980  
33981 34020  
34021 34060  
34061 34100  
34101 34140  
34141 34180  
34181 34220  
34221 34260  
34261 34300  
34301 34340  
34341 34380  
34381 34420  
34421 34460  
34461 34500  
34501 34540  
34541 34580  
34581 34620  
34621 34660  
34661

```
FT DOMAIN 3592 3632 LDL-RECEPTOR CLASS A 28.
FT DOMAIN 3633 3673 LDL-RECEPTOR CLASS A 29.
FT DOMAIN 3676 3716 LDL-RECEPTOR CLASS A 30.
FT DOMAIN 3717 3756 LDL-RECEPTOR CLASS A 31.
FT DOMAIN 3757 3795 LDL-RECEPTOR CLASS A 32.
FT DOMAIN 3796 3834 LDL-RECEPTOR CLASS A 33.
FT DOMAIN 3840 3880 LDL-RECEPTOR CLASS A 34.
FT DOMAIN 3881 3922 LDL-RECEPTOR CLASS A 35.
FT DOMAIN 3926 3964 LDL-RECEPTOR CLASS A 36.

Query Match 17.8%; Score 86.5; DB 1; Length 4655;
Best Local Similarity 26.1%; Pred. No. 2.1;
Matches 23; Conservative 14; Mismatches 30; Indels 21; Gaps 5;

Qy 6 CGENEKDYSCSKC-----DKKC-----KYDGVVEEDDEBNPCLVRVCHQD 49
Db 3745 CGDNDSEENACPRECTSEBPCVQIQIPSCWICDHYNDCGNSDER---DCENRTCHPE 3801

Qy 50 CV-CEEGFYRNKDKCVSAEDCELDNMD 76
Db 3802 YFQTSCHGVHSELKCDGSADC-LDASD 3828

RESULT 11
ICE2_ASCSU STANDARD; PRT; 65 AA.
AC P07852;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chymotrypsin/elastase isoforms 2 to 5.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoidea).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=84255715; PubMed=6564898;
RA Babin D.R., Feanasky R.J., Goss S.M.;
RT "The isoforms of chymotrypsin/elastase from Ascaris lumbricoidea: the primary structure."
RL Arch. Biochem. Biophys. 232:143-161(1984).
CC -!- FUNCTION: Defend the organism against the host's proteinases.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 TIL (trypsin inhibitory-like) domain.
DR HSP; P07851; IEAI.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
KW Serine protease inhibitor.
FT DOMAIN 4 59 TIL.
FT DISULFID 4 37 BY SIMILARITY.
FT DISULFID 13 32 BY SIMILARITY.
FT DISULFID 16 28 BY SIMILARITY.
FT DISULFID 20 59 BY SIMILARITY.
FT DISULFID 39 53 BY SIMILARITY.
FT ACT_SITE 30 31 REACTIVE BOND.
FT VARIANT 25 25 K -> N (IN INHIBITOR 2 AND 4).
FT VARIANT 40 40 T -> S (IN INHIBITOR 2 AND 4).
FT VARIANT 64 65 MISSING (IN INHIBITOR 2).
FT VARIANT 65 65 K -> E (IN INHIBITOR 3).
FT VARIANT 65 65 K -> R.
SQ SEQUENCE 65 AA; 7241 MW; B4E51CA166EA4BE3 CRC64;

Query Match 17.6%; Score 85.5; DB 1; Length 65;
Best Local Similarity 28.4%; Pred. No. 0.037;
Matches 19; Conservative 7; Mismatches 28; Indels 13; Gaps 3;

Qy 6 CGENEKDYSCSKCKKCKYDGVVEEDDEBNPCLVRVCHQDVCE--EGFYRNKDDK 63
Db 4 CGKNEVWTECTG--CELKC-----GQDEKTPCALMCRPPSCCECTPGRMRTHDGK 52

Qy 64 CVSAEDC 70
Db 64 CVSAEDC 70

Db 53 CVPVSEC 59

RESULT 12
LDVR_CHICK STANDARD; PRT; 863 AA.
ID LDVR_CHICK
AC P98165;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor)
DE (Vitellogenin receptor) (VTG receptor).
GN VLDLR OR VTGR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Ovary;
RX MEDLINE=95045409; PubMed=7957081;
RA Bujo H., Hermann M., Kaderli M.O., Jacobsen L., Sugawara S.,
RA Nimpf J., Yamamoto T., Schneider W.J.;
RT "Chicken oocyte growth is mediated by an eight ligand binding repeat member of the LDL receptor family."
RT member of the LDL receptor family."
RL EMBO J. 13:5165-5175(1994).
RN [2]
RP SEQUENCE OF 510-518; 546-554 AND 819-827.
RC STRAIN=White leghorn; TISSUE=Follicle membrane;
RX MEDLINE=92011638; PubMed=1655760;
RA Barber D.L., Sanders E.J., Aebersold R., Schneider W.J.;
RT "The receptor for yolk lipoprotein deposition in the chicken oocyte."
RL J. Biol. Chem. 266:18761-18770(1991).
CC -!- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must first cluster into clathrin-coated pits. Binding to Reelin induces tyrosine phosphorylation of Dab1 and modulation of Tau phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the extracellular matrix protein Reelin (By similarity).
CC -!- TISSUE SPECIFICITY: ABUNDANT IN OOCYTES; MUCH LESS IN HEART AND SKELETAL MUSCLE.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR EMBL; X80207; CAA56505.1; -.
DR PIR; S51789; S51789.
DR HSP; P01130; IAJJ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 8.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
```



QY 4 MCGGENEKYDCGSKEC-DKKCKYDGVBEEDDEPNVCLVRVCHQDCV-CEEGFYRNKD 61





```
FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
FT DISULFID 326 344 BY SIMILARITY.
FT DISULFID 338 355 BY SIMILARITY.
FT DISULFID 360 371 BY SIMILARITY.
FT DISULFID 367 380 BY SIMILARITY.
FT DISULFID 382 394 BY SIMILARITY.
FT DISULFID 400 410 BY SIMILARITY.
FT DISULFID 406 419 BY SIMILARITY.
FT DISULFID 421 434 BY SIMILARITY.
FT DISULFID 706 719 BY SIMILARITY.
FT DISULFID 715 734 BY SIMILARITY.
FT DISULFID 736 749 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 161 161 C -> G (IN REF. 1).
FT CONFLICT 262 262 P -> L (IN REF. 3).
FT CONFLICT 297 297 C -> S (IN REF. 2).
SQ SEQUENCE 873 AA; 96372 MW; 08F09F93825195CB CRC64;

Query Match 17.0%; Score 82.5; DB 1; Length 873;
Best Local Similarity 31.7%; Pred. No. 0.96;
Matches 26; Conservative 7; Mismatches 38; Indels 11; Gaps 4;

Qy 5 QCGENEKYDS-----CGSKEC-DKKCKYGVVEEDDEPNVPCLVRYCHOD-CVCEE 54
Db 228 QCGRPVHTKCTSETSCGSGEICHKKRCDGPDCKGSDGDEVNCPSTRCPDQFCEC 287

Qy 55 GFYRNKDDKCVSARDCELDNMD 76
Db 288 GSCIHSRQCNGIRDG-VDGSD 308

RESULT 15
ID LVDR RABIT STANDARD; PRT; 873 AA.
AC P35953;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor).
GN VLDLR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9302842; PubMed=1384047;
RA Takahashi S., Kawarabayashi Y., Nakai T., Sakai J., Yamamoto T.;
RT "Rabbit very low density lipoprotein receptor: a low density
RT lipoprotein receptor-like protein with distinct ligand specificity.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:9252-9256(1992).
CC -!- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must
CC first cluster into clathrin-coated pits. Binding to Reelin induces
CC tyrosine phosphorylation of Dab1 and modulation of tau
CC phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the extracellular matrix protein Reelin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN HEART, MUSCLE, AND ADIPOSE
CC TISSUE.
CC -!- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D11100; BAA01874.1; -.
DR PIR; A46286; QRRBYD.
DR HSSP; P01130; 1AJJ.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000333; LDL_receptor_rep.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00057; ldl_recept_a; 8.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 5.
DR PROSITE; PS00010; ASX HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 8.
DR PROSITE; PS00068; LDLRA_2; 8.
KW Glycoprotein; VLDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 873 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
FT DOMAIN 28 797 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 798 819 POTENTIAL.
FT DOMAIN 820 873 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 69 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 70 110 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 111 151 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 152 190 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 191 231 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 237 275 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 276 314 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 316 355 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 356 395 EGF-LIKE 1.
FT DOMAIN 396 435 EGF-LIKE 2.
FT REPEAT 439 480 CALCIUM-BINDING (POTENTIAL).
FT REPEAT 481 524 LDL-RECEPTOR CLASS B 1.
FT REPEAT 525 567 LDL-RECEPTOR CLASS B 2.
FT REPEAT 568 611 LDL-RECEPTOR CLASS B 3.
FT REPEAT 612 654 LDL-RECEPTOR CLASS B 4.
FT REPEAT 655 696 LDL-RECEPTOR CLASS B 5.
FT DOMAIN 702 750 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 751 790 EGF-LIKE 3.
FT SITE 832 837 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT DISULFID 33 45 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 40 58 BY SIMILARITY.
FT DISULFID 52 67 BY SIMILARITY.
FT DISULFID 72 84 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 91 108 BY SIMILARITY.
FT DISULFID 113 127 BY SIMILARITY.
FT DISULFID 120 140 BY SIMILARITY.
FT DISULFID 134 149 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 161 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 200 218 BY SIMILARITY.
FT DISULFID 212 229 BY SIMILARITY.
FT DISULFID 239 251 BY SIMILARITY.
FT DISULFID 246 264 BY SIMILARITY.
FT DISULFID 258 273 BY SIMILARITY.
FT DISULFID 278 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
```

Search completed: December 8, 2003, 09:14:12  
Job time : 24.0606 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 88.2424 Seconds  
(without alignments)  
245.646 Million cell updates/sec

Title: US-09-498-556C-59  
Perfect score: 486  
Sequence: 1 KATMQCENEKYDCSGKCE.....VSAEDCELDNMDFIYFGTRN 84

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTEMBL 23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rhodet.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	91	5 Q16938	Q16938 ancylostoma
2	196.5	40.4	102	5 Q962V8	Q962V8 ancylostoma
3	158.5	32.6	76	5 Q909B1	Q909B1 ancylostoma
4	139.5	28.7	98	5 Q16939	Q16939 ancylostoma
5	137.5	28.3	77	5 Q16935	Q16935 ancylostoma
6	137.5	28.3	91	5 Q16947	Q16947 ancylostoma
7	137.5	28.3	100	5 Q16940	Q16940 ancylostoma
8	125	25.7	432	5 Q814B8	Q814B8 caenorhabdi
9	125	25.7	537	5 Q901T6	Q901T6 caenorhabdi
10	124	25.5	102	12 Q9DQD5	Q9DQD5 microplitis
11	123.5	25.4	84	5 Q816Q5	Q816Q5 anopheles s
12	120.5	24.8	499	11 Q88714	Q88714 mus musculu
13	119.5	24.6	154	5 Q81TP8	Q81TP8 oesophagost
14	117	24.1	77	5 Q8T0W0	Q8T0W0 pimpha hypo
15	113	23.3	222	5 Q16620	Q16620 caenorhabdi
16	112	23.0	249	5 Q45764	Q45764 caenorhabdi

17	111.5	22.9	69	5 Q8MX86	Q8MX86 glossina mo
18	111	22.8	313	11 Q8K3U2	Q8K3U2 mus musculu
19	110	22.6	796	5 Q9UL75	Q9UL75 caenorhabdi
20	109.5	22.5	1642	5 Q62055	Q62055 caenorhabdi
21	106	21.8	731	5 Q814B9	Q814B9 caenorhabdi
22	105.5	21.7	145	5 Q18156	Q18156 caenorhabdi
23	100	20.6	5374	11 Q99ND0	Q99ND0 mus musculu
24	99	20.4	1036	5 Q97378	Q97378 strongyloce
25	98.5	20.3	86	5 Q817M1	Q817M1 caenorhabdi
26	98.5	20.3	135	5 Q18158	Q18158 caenorhabdi
27	98	20.2	137	5 Q18157	Q18157 caenorhabdi
28	97	20.0	3843	5 Q9VU94	Q9VU94 drosophila
29	97	20.0	3843	5 Q9U5D0	Q9U5D0 drosophila
30	96.5	19.9	869	13 Q42126	Q42126 xenopus lae
31	96	19.8	226	12 Q98332	Q98332 microplitis
32	96	19.8	315	12 Q98331	Q98331 microplitis
33	95.5	19.7	166	5 Q16488	Q16488 caenorhabdi
34	94.5	19.4	98	5 Q94162	Q94162 caenorhabdi
35	93.5	19.2	2108	13 Q98U19	Q98U19 gallus gall
36	92.5	19.0	2551	4 Q8WQ8	Q8WQ8 homo sapien
37	92.5	19.0	2551	4 Q8IUG9	Q8IUG9 homo sapien
38	91.5	18.8	62	5 Q77419	Q77419 ascaris suu
39	90.5	18.6	195	5 Q18805	Q18805 caenorhabdi
40	89.5	18.4	2559	11 Q9R4U0	Q9R4U0 mus musculu
41	87.5	18.0	949	5 P90956	P90956 caenorhabdi
42	87	17.9	92	5 Q9GPC4	Q9GPC4 drosophila
43	86.5	17.8	807	5 Q8MP01	Q8MP01 halocynthia
44	86	17.7	457	5 Q9NC91	Q9NC91 strongyloce
45	85	17.5	461	11 P97883	P97883 rattus norv

## ALIGNMENTS

### RESULT 1

Q16938	Q16938	PRELIMINARY;	PRT;	91 AA.
AC	Q16938;			
DT	01-NOV-1996 (Tremblrel. 01, Created)			
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	Anti-coagulant protein C2 precursor (fragment)			
OS	Ancylostoma caninum (Dog hookworm)			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;			
OC	Ancylostomatidae; Ancylostomatinae; Ancylostominae; Ancylostoma.			
OX	NCBI_TaxID=29170;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98298519; PubMed=9634780;			
RA	Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,			
RA	van den Brande I., Gansemans Y.G., Lauwereys M.J., Vlasuk G.P.,			
RA	Stanssens P.E.;			
RT	"surface expression and ligand-based selection of cDNAs fused to			
RT	filamentous phage gene VI."			
RL	Biotechnology 13:378-382(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96312555; PubMed=8700900;			
RA	Stanssens P., Bergum P.W., Gansemans Y., Jaspers L., Laroche Y.,			
RA	Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,			
RA	Lasters I., Vlasuk G.P.;			
RT	"Anticoagulant repertoire of the hookworm Ancylostoma caninum."			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).			
DR	EMBL; U30793; AAC47080.1; -			
DR	HSSP; P56682; ICCV.			
DR	InterPro; IPR006209; EGF like.			
DR	InterPro; IPR002919; TIL_Cysrich.			
DR	Pfam; PF01826; TIL; 1.			
DR	PROSITE; PS01186; EGF_2; 1.			
KW	Signal.			
FT	NON_TER	1	1	POTENTIAL.
FT	SIGNAL	<1	7	
FT	CHAIN	8	91	ANTI-COAGULANT PROTEIN C2.





DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Anticoagulant peptide (Fragment)  
OS Ancylostoma caninum (Dog hookworm)  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
OX NCBI\_TaxID=29170;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93274118; PubMed=8501344;  
RA Cappello M., Clyne L.P., McPhedran P., Hotez P.J.;  
RT "Ancylostoma factor Xa inhibitor: partial purification and its  
RT identification as a major hookworm-derived anticoagulant in vitro";  
RL J. Infect. Dis. 167:1474-1477(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97039675; PubMed=8885227;  
RA Cappello M., Hawdon J.M., Jones B.F., Kennedy W.P., Hotez P.J.;  
RT "Ancylostoma caninum anticoagulant peptide: cloning by PCR and  
RT expression of soluble, active protein in E. coli";  
RL Mol. Biochem. Parasitol. 80:113-117(1996).  
DR EMBL; U18305; AAC47318.1; -  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002919; TIL\_Cysrich.  
DR Pfam; PF01826; TIL; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 77 AA; 8746 MW; 7F9294472EEB1119 CRC64;  
Query Match 28.3%; Score 137.5; DB 5; Length 77;  
Best Local Similarity 42.5%; Pred. No. 9.2e-09;  
Matches 31; Conservative 11; Mismatches 26; Indels 5; Gaps 3;  
QY 1 KATWCCGENEYKDSGCS-KECDKKCKYDGVVEEDDEEENPVCLVRVCHQDCVCEEGFYRN 59  
DB 1 KAYPECGENWLDGCGTKPCAEKCNPEEPEDPICRSRGCLLPPA---CVCKDGFYRD 57  
QY 60 K-DDKCVSAEDCE 71  
DB 58 TWIGDCVREEDC 70  
RESULT 6  
ID Q16947 PRELIMINARY; PRT; 91 AA.  
AC Q16947;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Anticoagulant peptide isoform 2.  
GN ACAP-2  
OS Ancylostoma caninum (Dog hookworm)  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
OX NCBI\_TaxID=29170;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95320231; PubMed=7597095;  
RA Cappello M., Vlasuk G.P., Bergum P.W., Huang S., Hotez P.J.;  
RT "Ancylostoma caninum anticoagulant peptide: a hookworm-derived  
RT inhibitor of human coagulation factor Xa";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:6152-6156(1995).  
DR EMBL; U46550; AA887584.1; -  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002919; TIL\_Cysrich.  
DR Pfam; PF01826; TIL; 1.

DR PROSITE; PS01186; EGF\_2; 1.  
SQ SEQUENCE 91 AA; 10324 MW; E627F11E89B3EB4A CRC64;  
Query Match 28.3%; Score 137.5; DB 5; Length 91;  
Best Local Similarity 42.5%; Pred. No. 1.1e-08;  
Matches 31; Conservative 11; Mismatches 26; Indels 5; Gaps 3;  
QY 1 KATWCCGENEYKDSGCS-KECDKKCKYDGVVEEDDEEENPVCLVRVCHQDCVCEEGFYRN 59  
DB 15 KAYPECGENWLDGCGTKPCAEKCNPEEPEDPICRSRGCLLPPA---CVCKDGFYRD 71  
QY 60 K-DDKCVSAEDCE 71  
DB 72 TWIGDCVREEDC 84  
RESULT 7  
ID Q16940 PRELIMINARY; PRT; 100 AA.  
AC Q16940;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Anti-coagulant protein 5 precursor.  
OS Ancylostoma caninum (Dog hookworm)  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
OX NCBI\_TaxID=29170;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96312555; PubMed=8700900;  
RA Stanssens P., Bergum P.W., Ganssens Y., Jaspers L., Laroche Y.,  
RA Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,  
RA Lasters I., Vlasuk G.P.;  
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).  
DR EMBL; U30795; AAC47082.1; -  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002919; TIL\_Cysrich.  
DR Pfam; PF01826; TIL; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW Signal.  
FT CHAIN 1 23 POTENTIAL.  
FT CHAIN 23 100 ANTICOAGULANT PROTEIN 5.  
SQ SEQUENCE 100 AA; 11359 MW; DE9CB76729719A42 CRC64;  
Query Match 28.3%; Score 137.5; DB 5; Length 100;  
Best Local Similarity 42.5%; Pred. No. 1.2e-08;  
Matches 31; Conservative 11; Mismatches 26; Indels 5; Gaps 3;  
QY 1 KATWCCGENEYKDSGCS-KECDKKCKYDGVVEEDDEEENPVCLVRVCHQDCVCEEGFYRN 59  
DB 24 KAYPECGENWLDGCGTKPCAEKCNPEEPEDPICRSRGCLLPPA---CVCKDGFYRD 80  
QY 60 K-DDKCVSAEDCE 71  
DB 81 TWIGDCVREEDC 93  
RESULT 8  
ID Q814B8 PRELIMINARY; PRT; 432 AA.  
AC Q814B8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein Y69H2.3d.  
GN Y69H2.3d.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]



DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE GASTRIC MUCIN-LIKE protein (Fragment).  
 GN GASTRIC MUCIN-LIKE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stomach;  
 RA Tomasetto C., Masson R., Wendling C., Lefebvre O., Chenard M.P.,  
 RA Rio M.C.;  
 RT "Identification of interactions between trefoil peptides and members  
 RT of the mucin protein family using the yeast two-hybrid system.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ010752; CAA09343.1; -;  
 DR HSSP; P56682; ICCV.  
 DR InterPro; IPR002919; TIL\_Cysrich.  
 DR InterPro; IPR001846; VWF\_D.  
 DR Pfam; PF01826; TIL; 1.  
 DR Pfam; PF00094; vwd; 1.  
 DR SMART; SM00216; VWD; 1.  
 FT NON\_TER 1  
 FT NON\_TER 499 499  
 SQ SEQUENCE 499 AA; 54190 MW; 04F89EF4F23EE61E CRC64;  
 Query Match 24.8%; Score 120.5; DB 11; Length 499;  
 Best Local Similarity 30.1%; Pred. No. 5.7e-06;  
 Matches 25; Conservative 8; Mismatches 35; Indels 15; Gaps 2;  
 QY 6 CGENEKDYSCGSKCECDK-----KCKYDGVVEEDDEPNVPLVRVCHQDCVCEGFGYR 58  
 DB 14 CPEPKTFQSCQSSEDEKFGAACAPTQMLAT-----GIDCVPTKESGGCVCPKGLYE 65  
 QY 59 NKDDKCVSAEDCELDNMDFIYPG 81  
 DB 66 NSDQGVPAEPCDYGVSYPG 88  
 RESULT 13  
 Q8ITP8 PRELIMINARY; PRT; 154 AA.  
 AC Q8ITP8;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Putative trypsin-like inhibitor protein precursor.  
 GN MCRP.  
 OS Oesophagostomum dentatum.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Strongyloidea; Chabertiidae; Oesophagostomum.  
 OX NCBI\_TaxID=61180;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boag P.R., Ranganathan S., Newton S.E., Gasser R.B.;  
 RT "Identification of a male-specific nematode protein with two trypsin  
 RT like inhibitor domains.";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF399936; AAN32637.1; -;  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 154 PUTATIVE TRYPSIN-LIKE INHIBITOR PROTEIN.  
 SQ SEQUENCE 154 AA; 16564 MW; A7F566E3957DA819 CRC64;  
 Query Match 24.6%; Score 119.5; DB 5; Length 154;  
 Best Local Similarity 36.9%; Pred. No. 2.4e-06;  
 Matches 24; Conservative 8; Mismatches 20; Indels 13; Gaps 3;  
 QY 6 CGENEKDYSCGSKCECDKCKYDGVVEEDDEPNVPLVRVCHQDCVCEGFGYRNDKCV 65  
 DB 26 CGENEYNPCGN-HCEDTCSF-----TRRGCIAMCGPAACVCKEGFYRNSAGKC- 73

QY 66 SAEDC 70  
 DB 74 -TKDC 77

## RESULT 14

Q8TOW0 PRELIMINARY; PRT; 77 AA.  
 AC Q8TOW0;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Cysteine-rich venom protein 6 precursor.  
 GN CVP6.  
 OS Pimpla hypochondriaca.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonoidea;  
 OC Ichneumonidae; Pimplinae; Pimpla.  
 OX NCBI\_TaxID=135724;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Venom secretory;  
 RA Parkinson N.M., Richards E.H., Conyers C., Smith I., Edwards J.P.;  
 RT "Analysis of venom constituents from the parasitoid wasp Pimpla  
 RT hypochondriaca and cloning of a cDNA encoding a venom protein.";  
 RL Insect Biochem. Mol. Biol. 0:0-0(0).  
 DR EMBL; AJ438997; CAD27742.1; -;  
 DR InterPro; IPR002919; TIL\_Cysrich.  
 DR Pfam; PF01826; TIL; 1.  
 KW Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 77 CYSTEINE-RICH VENOM PROTEIN 6.  
 SQ SEQUENCE 77 AA; 8184 MW; 6889CB5F40D63DD6 CRC64;  
 Query Match 24.1%; Score 117; DB 5; Length 77;  
 Best Local Similarity 35.4%; Pred. No. 2.3e-06;  
 Matches 23; Conservative 8; Mismatches 22; Indels 12; Gaps 3;

QY 6 CGENEKDYSCGSKCECDKCKYDGVVEEDDEPNVPLVRVCHQDCVCEGFGYRNDKCV 65  
 DB 24 CGPNRVYKSCGT-GCPETC-----ENPDPC-DRACHQGCFCCKGLIQIDGNCI 71

## RESULT 15

O16620 PRELIMINARY; PRT; 222 AA.  
 AC O16620;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical protein F36H9.4.  
 GN F36H9.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Waterston R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R., Dante M., Kramer J., Twyman B.;

RT "The sequence of *C. elegans* cosmid F36H9.";   
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.   
 RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF016668; AA012440.1; -.

KW Hypothetical protein.

SQ SEQUENCE 222 AA; 25752 MW; A9A999F0D2EEA03F CRC64;

### Query Match

23.3%; Score 113; DB 5; Length 222;

### Best Local Similarity

9; Mismatches 18; Indels

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

QY 2 ATMQGENEKYDSCGSKECDKKCKYDGVVEEDDEEPNV-----PCLVRVCHQDCVCEEGF 56

Db 20 APKKCGPNDFKECGT-ACEANC-----AEGHVMFCTMQCIVNVCO-----CKDGF 64

QY 57 YRNKDDKCVSAE 68

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Search completed: December 8, 2003, 09:18:37

Job time : 90.2424 secv

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 74.6667 Seconds  
(without alignments)  
209.232 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1: KATMCGENKDYSCSKEC.....VSAEDCLDNMDIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	11	US-09-498-272-59
2	486	100.0	91	11	US-09-498-272-128
3	212	43.6	82	11	US-09-498-272-25
4	212	43.6	82	11	US-09-498-272-49
5	212	43.6	171	11	US-09-498-272-62
6	202	41.6	86	11	US-09-498-272-46
7	201	41.4	89	11	US-09-498-272-48
8	201	41.4	108	11	US-09-498-272-24
9	199	40.9	86	11	US-09-498-272-45
10	197.5	40.6	88	11	US-09-498-272-43
11	195.5	40.2	84	11	US-09-498-272-50
12	195.5	40.2	162	11	US-09-498-272-63
13	193.5	39.8	84	11	US-09-498-272-51
14	193.5	39.8	162	11	US-09-498-272-64
15	186	38.3	83	11	US-09-498-272-52

16	186	38.3	102	11	US-09-498-272-37	Sequence 27, Appl
17	186	38.3	161	11	US-09-498-272-65	Sequence 65, Appl
18	184.5	38.0	87	11	US-09-498-272-44	Sequence 44, Appl
19	178	36.6	78	11	US-09-498-272-47	Sequence 47, Appl
20	159	32.7	77	11	US-09-498-272-58	Sequence 58, Appl
21	159	32.7	96	11	US-09-498-272-23	Sequence 23, Appl
22	139.5	28.7	75	11	US-09-498-272-6	Sequence 6, Appl
23	139.5	28.7	75	11	US-09-498-272-41	Sequence 41, Appl
24	139.5	28.7	79	11	US-09-498-272-8	Sequence 8, Appl
25	139.5	28.7	98	11	US-09-498-272-21	Sequence 21, Appl
26	137.5	28.3	77	11	US-09-498-272-4	Sequence 4, Appl
27	137.5	28.3	77	11	US-09-498-272-40	Sequence 40, Appl
28	137.5	28.3	78	11	US-09-498-272-28	Sequence 28, Appl
29	137.5	28.3	78	11	US-09-498-272-56	Sequence 56, Appl
30	137.5	28.3	81	11	US-09-498-272-7	Sequence 7, Appl
31	137.5	28.3	100	11	US-09-498-272-20	Sequence 20, Appl
32	137	28.2	75	11	US-09-498-272-57	Sequence 57, Appl
33	137	28.2	94	11	US-09-498-272-22	Sequence 22, Appl
34	136.5	28.1	78	11	US-09-498-272-53	Sequence 53, Appl
35	136	28.0	75	11	US-09-498-272-26	Sequence 26, Appl
36	134.5	27.7	77	11	US-09-498-272-55	Sequence 55, Appl
37	134.5	27.7	78	11	US-09-498-272-54	Sequence 54, Appl
38	118	24.3	67	12	US-10-087-887-40	Sequence 40, Appl
39	111	22.8	74	11	US-09-498-272-42	Sequence 42, Appl
40	100	20.6	69	12	US-10-174-151-1	Sequence 1, Appl
41	100	20.6	69	12	US-10-174-151-2	Sequence 2, Appl
42	100	20.6	71	12	US-10-174-151-3	Sequence 3, Appl
43	100	20.6	73	12	US-10-174-151-4	Sequence 4, Appl
44	87	17.9	115	12	US-10-114-774-4	Sequence 4, Appl
45	84.5	17.4	58	11	US-09-498-272-60	Sequence 60, Appl

## ALIGNMENTS

### RESULT 1

US-09-498-272-59  
; Sequence 59, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Laureweys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Ganssema, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS: 356

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

```
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-498-272-59

Query Match 100.0%; Score 486; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 8.6e-42;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 2
US-09-498-272-128
; Sequence 128, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

```
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-09-498-272-128

Query Match 100.0%; Score 486; DB 11; Length 91;
Best Local Similarity 100.0%; Pred. No. 9.4e-42;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 8 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 67

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 3
US-09-498-272-25
; Sequence 25, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
```

/ /  
/ COUNTRY: U.S.A.  
/ ZIP: 90071  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: IBM P.C. DOS 5.0  
/ SOFTWARE: Word Perfect 5.1  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/498,272  
/ FILING DATE: 04-Feb-2000  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US95/13231  
/ FILING DATE: October 17, 1995  
/ APPLICATION NUMBER: 08/486,399  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/486,397  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/465,380  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/461,965  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/326,110  
/ FILING DATE: October 18, 1994  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: BIGGS, SUZANNE L.  
/ REGISTRATION NUMBER: 30,158  
/ REFERENCE/DOCKET NUMBER: 216/270  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (213) 489-1600  
/ TELEFAX: (213) 955-0440  
/ TELEX: 67-3510  
/ INFORMATION FOR SEQ ID NO: 25:  
/ SEQUENCE CHARACTERISTICS:  
/ TYPE: amino acid  
/ LENGTH: 82 amino acids  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ ORIGINAL SOURCE:  
/ ORGANISM: Ancylostoma ceylanicum  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-498-272-25

Query Match 43.6%; Score 212; DB 11; Length 82;  
Best Local Similarity 50.0%; Pred. No. 3e-14;  
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;  
QY 6 CGENEKYDCGS-KECDKKCKYGVVEEDDEBNVPCLVRVCHQD--CVCBEGFYRNKDD 62  
Db 4 CGSNERYSDCGNDKQCKERKCNEDDYKGD-----ACRSHVCERPACVCEGDFYRNKKG 58

QY 63 KCVSAEDCELDNMDFI 78  
Db 59 SCVESDDCEYDNMDFI 74

RESULT 4  
US-09-498-272-49  
/ Sequence 49, Application US/09498272  
/ Publication No. US20030113890A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Vlasuk, George Phillip  
/ Stanssens, Patrick Eric Hugo  
/ Messens, Joris Hilda Lieven  
/ Larocche, Yves Rene  
/ Jespers, Laurent Stephane  
/ Gansmans, Yannick Georges Jozef  
/ Moyle, Matthew  
/ Bergum, Peter W.  
/ TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
/ INHIBITORS AND ANTICOAGULANT  
/ PROTEIN

/ /  
/ NUMBER OF SEQUENCES: 356  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Lyon & Lyon  
/ STREET: 633 West Fifth Street  
/ Suite 4700  
/ CITY: Los Angeles  
/ STATE: California  
/ COUNTRY: U.S.A.  
/ ZIP: 90071  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: IBM P.C. DOS 5.0  
/ SOFTWARE: Word Perfect 5.1  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/498,272  
/ FILING DATE: 04-Feb-2000  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US95/13231  
/ FILING DATE: October 17, 1995  
/ APPLICATION NUMBER: 08/486,399  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/486,397  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/465,380  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/461,965  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/326,110  
/ FILING DATE: October 18, 1994  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: BIGGS, SUZANNE L.  
/ REGISTRATION NUMBER: 30,158  
/ REFERENCE/DOCKET NUMBER: 216/270  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (213) 489-1600  
/ TELEFAX: (213) 955-0440  
/ TELEX: 67-3510  
/ INFORMATION FOR SEQ ID NO: 49:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 82 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ ORIGINAL SOURCE:  
/ ORGANISM: Ancylostoma ceylanicum  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 49:  
US-09-498-272-49  
Query Match 43.6%; Score 212; DB 11; Length 82;  
Best Local Similarity 50.0%; Pred. No. 3e-14;  
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;  
QY 6 CGENEKYDCGS-KECDKKCKYGVVEEDDEBNVPCLVRVCHQD--CVCBEGFYRNKDD 62  
Db 4 CGSNERYSDCGNDKQCKERKCNEDDYKGD-----ACRSHVCERPACVCEGDFYRNKKG 58  
QY 63 KCVSAEDCELDNMDFI 78  
Db 59 SCVESDDCEYDNMDFI 74  
RESULT 5  
US-09-498-272-62  
/ Sequence 62, Application US/09498272  
/ Publication No. US20030113890A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Vlasuk, George Phillip  
/ Stanssens, Patrick Eric Hugo  
/ Messens, Joris Hilda Lieven  
/ Lauwereys, Marc Josef  
/ Larocche, Yves Rene

```
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
;
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
;
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
;
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-09-498-272-62
Query Match 43.6%; Score 212; DB 11; Length 171;
Best Local Similarity 50.0%; Pred. No. 6.4e-14;
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;
Qy 6 CGENEKDYSCGSKCKKCKYDGVVEEDDEPNVPCLVRYCHQD--CVCEEGFYRNKDD 62
Db 93 CGSNERYSDCGNDKQCKRCNEDDYKGDG-----ACKSHVCVRPGACVCDGYRNKKG 147
Qy 63 KCVSAEDCELDNMDPI 78
Db 148 SCVESDDCEYDNMDPI 163
RESULT 6
US-09-498-272-46
; Sequence 46, Application US/09498272
```

```
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlausk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
;
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
;
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
;
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-498-272-46
Query Match 41.6%; Score 202; DB 11; Length 86;
Best Local Similarity 47.4%; Pred. No. 3.2e-13;
Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;
Qy 5 QCGENEKDYSCGSKCKKCKYDGVVEEDD--PEPNVPCLVRYCHQ--DCVCEEGFYRNK 60
Db 5 KCGPGERLDCANKKPCPKCIETSEEDDDVETDVRCLRVCPKLCICKDGYRNK 64
Qy 61 DDKCVSAEDCELDNMDPI 78
```



Query Match 41.4%; Score 201; DB 11; Length 89;  
Best Local Similarity 48.1%; Pred. No. 4.2e-13;

```
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-498-272-24

Query Match      41.4%; Score 201; DB 11; Length 108;
Best Local Similarity 48.1%; Pred. No. 5.1e-13;
Matches 36; Conservative 13; Mismatches 24; Indels 4; Gaps 3;

Qy 3 TMOQGENEKYDSCGS-KECDKKCKYDGVBEEDDEPNVPCLVRVCHQD--CVCEBGFYRN 59
Db 26 TNACGLNEYFAECNMGKECHRCN-EEENEERDEERITACLRVCPFGACVCKDGYRN 84

Qy 60 KDKCVSAEDCELDNMF1 78
Db 85 RTGSCVBEEDDCEYENMEFI 103

RESULT 9
US-09-498-272-45
; Sequence 45, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlaauk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieveen
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440

; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-498-272-45

Query Match      40.9%; Score 199; DB 11; Length 86;
Best Local Similarity 46.2%; Pred. No. 6.4e-13;
Matches 36; Conservative 16; Mismatches 22; Indels 4; Gaps 2;

Qy 5 QCGENEKDYSCGSKECDKKCKYDGVBEEDD--EENVPCLVRVCHQ--DCVCEBGFYRN 60
Db 5 KCGPGERLDCANKPCPCPKIETSEEDDDVEDTDVRLVRCERPLKCKICKDGYRNK 64

Qy 61 DDKCVSAEDCELDNMF1 78
Db 65 KGEVTDVDCQEDFMEFI 82

RESULT 10
US-09-498-272-43
; Sequence 43, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlaauk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieveen
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
```



```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-09-498-272-63

Query Match 40.2%; Score 195.5; DB 11; Length 162;
Best Local Similarity 50.0%; Pred. No. 2.8e-12;
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;

Qy 1 KATWQGENEKYDSCGS-KECDKKCKYDGVVEEDDEPNVPLRVVC-HQDCVCEEGFYR 58
Db 1 KSAKKCGLNKLD-CGNLKACEKKCSLDLDNEEDYGEEDSKRSRRCVVCDEGFYR 59

Qy 59 NKDDKCVSAEDCELDNMDFI 78
Db 60 NKGQCVRDCEYDNMEII 79

RESULT 13
US-09-498-272-51
; Sequence 51, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlaeuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
```

```
;
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-498-272-51

Query Match 39.8%; Score 193.5; DB 11; Length 84;
Best Local Similarity 50.0%; Pred. No. 2.2e-12;
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;

Qy 1 KATWQGENEKYDSCGS-KECDKKCKYDGVVEEDDEPNVPLRVVC-HQDCVCEEGFYR 58
Db 1 KSAKKCGLNKLD-CGNLKACEKKCSLDLDNEEDYGEEDSKRSRRCVVCDEGFYR 59

Qy 59 NKDDKCVSAEDCELDNMDFI 78
Db 60 NKGQCVRDCEYDNMEII 79

RESULT 14
US-09-498-272-64
; Sequence 64, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlaeuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
```

STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: California  
ZIP: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-09-498-272-64

Query Match 39.8%; Score 193.5; DB 11; Length 162;  
Best Local Similarity 50.0%; Pred. No. 4.4e-12;  
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;  
QY 1 KATMCGENKDYDCGS-KECDKKCKYDGVVEEDDEPNVCLVRVC-HQDCVCEGFYR 58  
Db 1 KSAKCKGLNEKLD-CGNLKACEKKCSLDNEEDYGEEDSKCRSEICGRVVCDEGFYR 59  
QY 59 NKDDKCVSAEDCELDNDFI 78  
Db 60 NKKGQCVTRDDCEYDNMEII 79

RESULT 15  
US-09-498-272-52  
Sequence 52, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansemans, Yannick Georges Jozef  
Moyle, Matthew

Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: California  
ZIP: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma duodenale  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
US-09-498-272-52

Query Match 38.3%; Score 186; DB 11; Length 83;  
Best Local Similarity 50.0%; Pred. No. 1.3e-11;  
Matches 40; Conservative 11; Mismatches 25; Indels 4; Gaps 4;  
QY 1 KATMCGENKDYDCGS-KECDKKCKYDGVVEEDDEPNVCLVRVC-HQDCVCEGFYR 58  
Db 1 KAAKCKGLNERLD-CGNLKQCEPKCS-DLESEVEEDESCKRSECSRRVVCVDEGFYR 58  
QY 59 NKDDKCVSAEDCELDNDFI 78  
Db 59 NKKGKCVAKDVCEDDNMEII 78

Search completed: December 8, 2003, 09:34:03  
Job time : 75.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 114.545 seconds  
(without alignments)  
116.399 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMOGENEYDSCGSKC.....VSAEDCLDNMDFIYFGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	20	AA1980.DAT.*
2	486	100.0	84	21	AA1981.DAT.*
3	486	100.0	91	17	AA1982.DAT.*
4	486	100.0	91	20	AA1983.DAT.*
5	486	100.0	91	20	AA1984.DAT.*
6	486	100.0	91	21	AA1985.DAT.*
7	212	43.6	82	20	AA1986.DAT.*
8	212	43.6	82	20	AA1987.DAT.*
9	212	43.6	82	21	AA1988.DAT.*

10	212	43.6	82	21	AA1989.DAT.*
11	212	43.6	171	20	AA1990.DAT.*
12	212	43.6	190	17	AA1991.DAT.*
13	212	43.6	190	20	AA1992.DAT.*
14	202	41.6	86	20	AA1993.DAT.*
15	202	41.6	86	21	AA1994.DAT.*
16	202	41.6	105	17	AA1995.DAT.*
17	202	41.6	105	20	AA1996.DAT.*
18	201	41.4	89	20	AA1997.DAT.*
19	201	41.4	108	20	AA1998.DAT.*
20	201	41.4	108	21	AA1999.DAT.*
21	199	40.9	86	20	AA2000.DAT.*
22	199	40.9	86	21	AA2001.DAT.*
23	199	40.9	105	17	AA2002.DAT.*
24	199	40.9	105	20	AA2003.DAT.*
25	197.5	40.6	88	20	AA2004.DAT.*
26	197.5	40.6	88	21	AA2005.DAT.*
27	197.5	40.6	107	17	AA2006.DAT.*
28	197.5	40.6	107	20	AA2007.DAT.*
29	195.5	40.2	84	20	AA2008.DAT.*
30	195.5	40.2	162	20	AA2009.DAT.*
31	195.5	40.2	162	21	AA2010.DAT.*
32	195.5	40.2	181	17	AA2011.DAT.*
33	195.5	40.2	181	20	AA2012.DAT.*
34	193.5	39.8	84	20	AA2013.DAT.*
35	193.5	39.8	162	20	AA2014.DAT.*
36	193.5	39.8	162	21	AA2015.DAT.*
37	193.5	39.8	181	17	AA2016.DAT.*
38	193.5	39.8	181	20	AA2017.DAT.*
39	186	38.3	83	20	AA2018.DAT.*
40	186	38.3	102	20	AA2019.DAT.*
41	186	38.3	102	21	AA2020.DAT.*
42	186	38.3	161	20	AA2021.DAT.*
43	186	38.3	180	17	AA2022.DAT.*
44	186	38.3	180	20	AA2023.DAT.*
45	184.5	38.0	87	20	AA2024.DAT.*

## ALIGNMENTS

RESULT 1

AA1980.DAT.\*

ID AA1980.DAT.\* standard; Protein; 84 AA.

AC AA1980.DAT.\*

XX AA1980.DAT.\*

DT 15-NOV-1999 (first entry)

DE Mature nematode extracted anticoagulant protein AcanAPC2.

DE Mature nematode extracted anticoagulant protein AcanAPC2.

KW Nematode extracted anticoagulant protein; NAP; anticoagulant;

KW serine protease inhibitor; NAP domain; factor VIIa/TF.

XX Ancylostoma caninum.

OS Ancylostoma caninum.

PN US5955294-A.

XX US5955294-A.

PD 21-SEP-1999.

XX 21-SEP-1999.

PF 19-APR-1996; 96US-0634641.

XX 19-APR-1996; 96US-0634641.

PR 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0461965.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

PA (CORV-) CORVAS INT INC.

XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

```
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH,
PI Vlasuk GP;
XX WPI; 1999-539569/45.
XX Screening an isolated protein for Nematode-extracted Anticoagulant
XX Protein domains
XX Disclosure; Columns 142-144; 197pp; English.
XX The present sequence represents a nematode extracted anticoagulant
XX protein (NAP). The protein has activity as an anticoagulant and/or serine
XX protease inhibitor. The protein contains at least one NAP domain which
XX has selective inhibitory activity for factor VIIa/TF. The specification
XX describes a method for screening an isolated protein at least one domain
XX for factor VIIa/TF selective inhibitory activity. The method comprises
XX determining the time to clotting effected by a concentration of the
XX isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
XX activated partial thromboplastin time (aPTT) assay; calculating
XX prolongation of clotting effected by the isolated protein in each of
XX the PT and aPTT assay, with respect to a baseline clotting value for
XX each assay, where prolongation of clotting is calculated as fold
XX elevation of clotting time relative to a baseline clotting value, where
XX a doubling of clotting time is deemed a two-fold elevation; and
XX calculating a PT to aPTT prolongation ratio, where a ratio at least
XX one is indicative of factor VIIa/TF inhibitory activity. The method is
XX useful for determining if a protein has factor VIIa/TF inhibitory
XX activity.
XX SQ Sequence 84 AA;
Query Match 100.0%; Score 486; DB 20; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATWQCGNEKYDSCGSKCKKCYDGVREEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
DB 1 KATWQCGNEKYDSCGSKCKKCYDGVREEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
DB 61 DDKCVSAEDCELDNMDFIYPGTRN 84
RESULT 2
AAB15317
ID AAB15317 standard; Protein; 84 AA.
AC AAB15317;
XX 19-DEC-2000 (first entry)
DE A. caninum nematode-extracted anticoagulant protein AcanAPc2(mature).
KW Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;
KW canine hookworm; thrombosis; vaccine.
XX Ancylostoma caninum.
OS US6087487-A.
PN 11-JUL-2000.
XX 12-FEB-1999; 99US-0249451.
XX 17-OCT-1995; 95WO-US13231.
PR 17-APR-1997; 97US-0809455.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
XX PA
PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
XX WPI; 2000-531359/48.
XX N-PSDB; AAA73373.
XX New cDNA molecule encoding a protein having factor Xa inhibitory
XX PT activity for preventing and treating blood clotting disorders,
XX comprises nematode-extracted anticoagulant protein domains -
XX Example A; Fig 16; 197pp; English.
XX The present sequence is the Ancylostoma caninum nematode-extracted
XX anticoagulant protein AcanAPc2. Proteins of this kind have been shown to
XX be effective at preventing blood clotting without causing excessive
XX bleeding. The protein can be used in blood collection tubes to aid the
XX isolation of plasma from the blood, to prevent thrombosis which may be
XX linked to the rupture of an atherosclerotic plaque, acute myocardial
XX infarction, angina, thrombolytic therapy, percutaneous transluminal
XX coronary angioplasty, disseminated intravascular coagulopathy, infection,
XX cancer and septic shock, and to produce antibodies. In the latter
XX instance, the antibodies can be raised in order to detect infection by
XX nematodes (the coding sequence can also be used for this) or as
XX diagnostic tests. The protein can also be used as a vaccine against
XX nematode parasites.
XX SQ Sequence 84 AA;
Query Match 100.0%; Score 486; DB 21; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATWQCGNEKYDSCGSKCKKCYDGVREEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
DB 1 KATWQCGNEKYDSCGSKCKKCYDGVREEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
DB 61 DDKCVSAEDCELDNMDFIYPGTRN 84
RESULT 3
AAR91701
ID AAR91701 standard; Protein; 91 AA.
AC AAR91701;
XX 25-MAR-2003 (updated)
DT 17-NOV-1996 (first entry)
XX AcanAPc2.
XX AcanAP; HpoNAP; NamNAP; AceNAP; AduNAP; anticoagulant;
KW nematode-extracted anticoagulant protein; serine protease;
KW nematode; thrombosis; parasitic worm.
XX Ancylostoma caninum.
OS WO9612021-A2.
PN 25-APR-1996.
XX 17-OCT-1995; 95WO-US13231.
XX 05-JUN-1995; 95US-0486399.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
XX PA
PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
XX WPI; 1999-539569/45.
XX Screening an isolated protein for Nematode-extracted Anticoagulant
XX Protein domains
XX Disclosure; Columns 142-144; 197pp; English.
XX The present sequence represents a nematode extracted anticoagulant
XX protein (NAP). The protein has activity as an anticoagulant and/or serine
XX protease inhibitor. The protein contains at least one NAP domain which
XX has selective inhibitory activity for factor VIIa/TF. The specification
XX describes a method for screening an isolated protein at least one domain
XX for factor VIIa/TF selective inhibitory activity. The method comprises
XX determining the time to clotting effected by a concentration of the
XX isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
XX activated partial thromboplastin time (aPTT) assay; calculating
XX prolongation of clotting effected by the isolated protein in each of
XX the PT and aPTT assay, with respect to a baseline clotting value for
XX each assay, where prolongation of clotting is calculated as fold
XX elevation of clotting time relative to a baseline clotting value, where
XX a doubling of clotting time is deemed a two-fold elevation; and
XX calculating a PT to aPTT prolongation ratio, where a ratio at least
XX one is indicative of factor VIIa/TF inhibitory activity. The method is
XX useful for determining if a protein has factor VIIa/TF inhibitory
XX activity.
XX SQ Sequence 84 AA;
Query Match 100.0%; Score 486; DB 20; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATWQCGNEKYDSCGSKCKKCYDGVREEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
DB 1 KATWQCGNEKYDSCGSKCKKCYDGVREEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
DB 61 DDKCVSAEDCELDNMDFIYPGTRN 84
RESULT 2
AAB15317
ID AAB15317 standard; Protein; 84 AA.
AC AAB15317;
XX 19-DEC-2000 (first entry)
DE A. caninum nematode-extracted anticoagulant protein AcanAPc2(mature).
KW Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;
KW canine hookworm; thrombosis; vaccine.
XX Ancylostoma caninum.
OS US6087487-A.
PN 11-JUL-2000.
XX 12-FEB-1999; 99US-0249451.
XX 17-OCT-1995; 95WO-US13231.
PR 17-APR-1997; 97US-0809455.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
XX PA
```



XX Vlasuk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;  
PI Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;  
XX WPI; 1996-222007/22.  
DR N-PSDB; AAT12947.  
XX  
PT Proteins with anticoagulant and/or serine protease inhibitory  
PT activity - isolated from nematodes and useful to inhibit blood  
PT coagulation  
XX  
PS Claim 89 + 96; Fig 9; 243pp; English.  
XX  
XX Proteins with anticoagulant and/or serine protease inhibitory  
CC activity, isolated from nematodes, are useful to inhibit blood  
CC coagulation. The proteins can be added to blood collection tubes  
CC defining the collection of mammalian plasma. They are also useful  
CC to prevent or inhibit thrombosis, and may be given alone or in  
CC combination with other therapeutic or in vivo diagnostic agents.  
CC The proteins can serve as immunogens to raise antibodies for use in  
CC the diagnosis and identification of NAP concn. levels in biological  
CC fluids, e.g. to detect mammalian infection with a parasitic worm.  
CC They can also be used as immunogens in prophylactic and therapeutic  
CC vaccines against parasitic worm infection. The proteins may  
CC double the clotting time of human plasma in prothrombin time assays  
CC when present at 10-50 nMol, and double the clotting time of human  
CC plasma in activated partial thrombin time assays when present  
CC at 10-100 nMol.  
CC The anticoagulant proteins are pref. derived from  
CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator  
CC americanus or Heligmosomoides polygyrus.  
CC The proteins pref. have 2 NAP domains and specifically inhibit  
CC the catalytic activity of the factor VIIa/TF complex in the  
CC presence of factor Xa or a catalytically inactive factor Xa deriv.,  
CC do not specifically inhibit the activation of factor VIIa in the  
CC absence of TF and do not specifically inhibit prothrombinase.  
CC (Updated on 25-MAR-2003 to correct PI field.)  
XX  
SQ Sequence 91 AA;  
Query Match 100.0%; Score 486; DB 17; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.9e-38;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 60  
Db 8 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 67  
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91  
RESULT 4  
AAY30393  
ID AAY30393 standard; Protein; 91 AA.  
XX AAY30393;  
XX  
XX 15-NOV-1999 (first entry)  
XX  
XX Nematode extracted anticoagulant protein AcanAPc2.  
XX  
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;  
KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
XX  
XX Ancylostoma caninum.  
OS  
XX US5955294-A.  
XX  
XX 21-SEP-1999.  
XX  
XX 19-APR-1996; 96US-0634641.  
XX

XX 19-APR-1996; 96US-0634641.  
PR 18-OCT-1994; 94US-0326110.  
PR 05-JUN-1995; 95US-0461965.  
PR 05-JUN-1995; 95US-0465380.  
PR 05-JUN-1995; 95US-0486397.  
PR 05-JUN-1995; 95US-0486399.  
PR 17-OCT-1995; 95WO-US13231.  
XX (CORV-) CORVAS INT INC.  
PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
XX Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
PI Vlasuk GP;  
XX WPI; 1999-539569/45.  
DR N-PSDB; AA299999.  
XX  
PT Screening an isolated protein for Nematode-extracted Anticoagulant  
PT Protein domains  
XX  
PS Example 13; Fig 9; 197pp; English.  
XX  
CC The present sequence represents a nematode extracted anticoagulant  
CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
CC protease inhibitor. The protein contains at least one NAP domain which  
CC has selective inhibitory activity for factor VIIa/TF. The specification  
CC describes a method for screening an isolated protein at least one domain  
CC for factor VIIa/TF selective inhibitory activity. The method comprises  
CC isolating a protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
CC activated partial thromboplastin time (APTT) assay; calculating  
CC the PT and APTT assay, with respect to a baseline clotting value for  
CC each assay, where prolongation of clotting is calculated as fold  
CC elevation of clotting time relative to a baseline clotting value, where  
CC a doubling of clotting time is deemed a two-fold elevation; and  
CC calculating a PT to APTT prolongation ratio, where a ratio at least  
CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
CC useful for determining if a protein has factor VIIa/TF inhibitory  
CC activity.  
XX  
SQ Sequence 91 AA;  
Query Match 100.0%; Score 486; DB 20; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.9e-38;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 60  
Db 8 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 67  
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91  
RESULT 5  
AAY30454  
ID AAY30454 standard; Protein; 91 AA.  
XX AAY30454;  
XX  
XX 15-NOV-1999 (first entry)  
XX  
XX Nematode extracted anticoagulant protein AcanAPc2.  
DE  
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;  
KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
XX  
XX Ancylostoma caninum.  
OS  
XX US5955294-A.  
XX

XX PD 21-SEP-1999.  
 XX PF 19-APR-1996; 96US-0634641.  
 XX PR 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX (CORV-) CORVAS INT INC.  
 XX PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
 XX PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX WPI; 1999-539569/45.  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains  
 DR Disclosure; Columns 175-176; 197pp; English.  
 XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of  
 CC the PT and APTT assay, with respect to a baseline clotting value for  
 CC each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where  
 CC a doubling of clotting time is deemed a two-fold elevation; and  
 CC calculating a PT to APTT prolongation ratio, where a ratio at least  
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
 CC useful for determining if a protein has factor VIIa/TF inhibitory  
 CC activity.  
 XX SQ Sequence 91 AA;  
 Query Match 100.0%; Score 486; DB 20; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-38;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KATMQCGENEKYDSCGSKCEDKKCKYDGVVEEDDEEPNVPCLVRVCHQDVCCEGFYRNK 60  
 DB 8 KATMQCGENEKYDSCGSKCEDKKCKYDGVVEEDDEEPNVPCLVRVCHQDVCCEGFYRNK 67  
 QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
 DB 68 DDKCVSAEDCELDNMDFIYPGTRN 91  
 RESULT 6  
 AAB15346  
 ID AAB15346 standard; Protein; 91 AA.  
 XX AAB15346;  
 AC AAB15346;  
 XX 19-DEC-2000 (first entry)  
 DT A. caninum nematode-extracted anticoagulant protein AcanAPc2.  
 DE Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;  
 KW Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;  
 KW canine hookworm; thrombosis; vaccine.  
 XX

OS Ancylostoma caninum.  
 XX US6087487-A.  
 PN 11-JUL-2000.  
 XX 12-FEB-1999; 99US-0249451.  
 XX 17-OCT-1995; 95WO-US13231.  
 PR 17-APR-1997; 97US-0809455.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 XX (CORV-) CORVAS INT INC.  
 XX PA Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 PI WPI; 2000-531359/48.  
 DR N-PSDB; AAA73373.  
 DR New cDNA molecule encoding a protein having factor Xa inhibitory  
 PT activity for preventing and treating blood clotting disorders,  
 PT comprises nematode-extracted anticoagulant protein domains -  
 PS Disclosure; Fig 11; 197pp; English.  
 XX The present sequence is the Ancylostoma caninum nematode-extracted  
 CC anticoagulant protein AcanAPc2. Proteins of this kind have been shown to  
 CC be effective at preventing blood clotting without causing excessive  
 CC bleeding. The protein can be used in blood collection tubes to aid the  
 CC isolation of plasma from the blood, to prevent thrombosis which may be  
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 CC cancer and septic shock, and to produce antibodies. In the latter  
 CC instance, the antibodies can be raised in order to detect infection by  
 CC nematodes (the coding sequence can also be used for this) or as  
 CC diagnostic tests. The protein can also be used as a vaccine against  
 CC nematode parasites.  
 XX SQ Sequence 91 AA;  
 Query Match 100.0%; Score 486; DB 21; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-38;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KATMQCGENEKYDSCGSKCEDKKCKYDGVVEEDDEEPNVPCLVRVCHQDVCCEGFYRNK 60  
 DB 8 KATMQCGENEKYDSCGSKCEDKKCKYDGVVEEDDEEPNVPCLVRVCHQDVCCEGFYRNK 67  
 QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
 DB 68 DDKCVSAEDCELDNMDFIYPGTRN 91  
 RESULT 7  
 AAY30399  
 ID AAY30399 standard; Protein; 82 AA.  
 XX AAY30399;  
 AC AAY30399;  
 XX 15-NOV-1999 (first entry)  
 DT Nematode extracted anticoagulant protein AcanAP4d2.  
 DE Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 XX serine protease inhibitor; NAP domain; factor VIIa/TF.  
 KW Ancylostoma ceylanicum.  
 XX





XX US955294-A.  
 XX 21-SEP-1999.  
 XX  
 XX 19-APR-1996; 96US-0634641.  
 XX  
 XX 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR,  
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PER,  
 PI Vlausuk GP;  
 XX WPI; 1999-539569/45.  
 XX  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains  
 XX  
 XX Disclosure; Fig 17; 197pp; English.  
 XX  
 XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (aPTT) assay; calculating  
 CC the PT and aPTT assay, with respect to the isolated protein in each of  
 CC each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where  
 CC a doubling of clotting time is deemed a two-fold elevation; and  
 CC calculating a PT to aPTT prolongation ratio, where a ratio at least  
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
 CC useful for determining if a protein has factor VIIa/TF inhibitory  
 CC activity.  
 XX  
 XX Sequence 171 AA;  
 SQ  
 Query Match 43.6%; Score 212; DB 20; Length 171;  
 Best Local Similarity 50.0%; Pred. No. 1.9e-12;  
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;  
 QY 6 CGENEKDYSCGS-KCEDKKYKDYGVVEEDDEBNVFLVRVCHQD--CVCSEGFYRNKDD 62  
 DB 93 CGSNERYSDCGNDKQCKERCNEDDYKGDGDE-----ACRSHVCERPACVCEGDFYRNKKG 147  
 QY 63 KCVSAEDCELDNNDFI 78  
 DB 148 SCVESDCCBYDNDNFI 163  
 RESULT 12  
 AAR91710  
 ID AAR91710 standard; Protein; 190 AA.  
 XX  
 XX AAR91710;  
 AC  
 DT 25-MAR-2003 (updated)  
 DT 17-NOV-1996 (first entry)  
 XX  
 XX AcenAP4.  
 DE  
 XX

KW AcanAP; HpoNAP; NamNAP; AceNAP; AdunAP; antioagulant;  
 KW nematode-extracted anticoagulant protein; serine protease;  
 KW nematode; thrombosis; parasitic worm.  
 XX  
 OS Ancylostoma ceylanicum.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..108  
 FT /label= AceNAP4d1  
 FT 109..190  
 FT /label= AceNAP4d2  
 XX  
 PN WO9612021-A2.  
 XX  
 PD 25-APR-1996.  
 XX  
 PF 17-OCT-1995; 95WO-US13231.  
 XX  
 PR 05-JUN-1995; 95US-0486399.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Vlausuk GP, Stanssens PER, Messens JHL, Lauwereys MJ, Laroche YR,  
 PI Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;  
 XX WPI; 1996-222007/22.  
 XX N-PSDB; AAT12956.  
 DR  
 XX Proteins with anticoagulant and/or serine protease inhibitory  
 PT activity - isolated from nematodes and useful to inhibit blood  
 PT coagulation  
 XX  
 XX Claim 221; Fig 7A; 243pp; English.  
 PS  
 CC Proteins with anticoagulant and/or serine protease inhibitory  
 CC activity, isolated from nematodes, are useful to inhibit blood  
 CC coagulation. The proteins can be added to blood collection tubes  
 CC defining the collection of mammalian plasma. They are also useful  
 CC to prevent or inhibit thrombosis, and may be given alone or in  
 CC combination with other therapeutic or in vivo diagnostic agents.  
 CC The proteins can serve as immunogens to raise antibodies for use in  
 CC the diagnosis and identification of NAP concn. levels in biological  
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.  
 CC They can also be used as immunogens in prophylactic and therapeutic  
 CC vaccines against parasitic worm infection. The proteins may  
 CC double the clotting time of human plasma in prothrombin time assays  
 CC when present at 10-50 nMol, and double the clotting time of human  
 CC plasma in activated partial thrombin time assays when present  
 CC at 10-100 nMol.  
 CC The anticoagulant proteins are pref. derived from  
 CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator  
 CC americanus or Heligmosomoides polygyrus  
 CC The proteins pref. have 2 NAP domains and specifically inhibit  
 CC the catalytic activity of the factor VIIa/TF complex in the  
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,  
 CC do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase.  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 XX Sequence 190 AA;  
 SQ

Query Match 43.6%; Score 212; DB 17; Length 190;  
 Best Local Similarity 50.0%; Pred. No. 2.1e-12;  
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENEKDYSCGS-KCEDKKYKDYGVVEEDDEBNVFLVRVCHQD--CVCSEGFYRNKDD 62  
 DB 112 CGSNERYSDCGNDKQCKERCNEDDYKGDGDE-----ACRSHVCERPACVCEGDFYRNKKG 166

QY 63 KCVSAEDCELDNMDFI 78  
 ||:|||||  
 Db 167 SCVESDDCEYDNMDFI 182

RESULT 13  
 AAY30384  
 ID AAY30384 standard; Protein; 190 AA.

AC AAY30384;  
 XX  
 DT 15-NOV-1999 (first entry)  
 XX  
 DE Nematode extracted anticoagulant protein AcenAP4.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Ancylostoma ceylanicum.

XX US9595294-A.

XX 21-SEP-1999.

XX 19-APR-1996; 96US-0634641.

XX 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;

PI Vlasuk GP;

XX WPI; 1999-539569/45.

DR NP-SDB; AAZ10441.

XX Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains

PS Example 9; Fig 7A; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (aPTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of  
 CC the PT and aPTT assay, with respect to a baseline clotting value for  
 CC each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where  
 CC a doubling of clotting time is deemed a two-fold elevation; and  
 CC calculating a PT to aPTT prolongation ratio, where a ratio at least  
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
 CC useful for determining if a protein has factor VIIa/TF inhibitory  
 CC activity.

XX Sequence 190 AA;

Query Match 43.6%; Score 212; DB 20; Length 190;

Best Local Similarity 50.0%; Pred. NO. 2.1e-12;

Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENKDYSGS-KCKDKKYDGVVEEDDEBNVPCLVRVCHOD--CVCEEGFYRNKDD 62  
 ||:|||||  
 Db 112 CGSNERYSDCGNDKQCKERKCNEDDYKGDG-----ACRSHVCERPACVCEDFYRNKKG 166

QY 63 KCVSAEDCELDNMDFI 78  
 ||:|||||  
 Db 167 SCVESDDCEYDNMDFI 182

RESULT 14

AAY30419

ID AAY30419 standard; Protein; 86 AA.

XX AAY30419;

AC AAY30419;

XX 15-NOV-1999 (first entry)

XX Mature nematode extracted anticoagulant protein AcanAP44.

XX Nematode extracted anticoagulant protein; NAP; anticoagulant;

XX serine protease inhibitor; NAP domain; factor VIIa/TF.

XX Ancylostoma caninum.

XX US9595294-A.

XX 21-SEP-1999.

XX 19-APR-1996; 96US-0634641.

XX 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;

PI Vlasuk GP;

XX WPI; 1999-539569/45.

XX Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains

PS Disclosure; Columns 131-134; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (aPTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of  
 CC the PT and aPTT assay, with respect to a baseline clotting value for  
 CC each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where  
 CC a doubling of clotting time is deemed a two-fold elevation; and  
 CC calculating a PT to aPTT prolongation ratio, where a ratio at least  
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
 CC useful for determining if a protein has factor VIIa/TF inhibitory  
 CC activity.

XX Sequence 86 AA;

Query Match 41.6%; Score 202; DB 20; Length 86;

Best Local Similarity 47.4%; Pred. No. 8.3e-12;  
Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;  
QY 5 OCGENEKYDSCGSKCKYDGVVEEDD--BEPNVPCLVRYCHQ--DCVCEEGFYRNK 60  
Db 5 KCGPGERLDCANKKPCPKIETSEEDDDVEETDVRCLVRVCERPLKICKDGYRNK 64  
QY 61 DDKCVSAEDCELDNMDPI 78  
Db 65 KGECVTDDVCQEDFMEFI 82

RESULT 15  
AAB15304  
ID AAB15304 standard; Protein; 86 AA.  
XX  
AC AAB15304;  
XX  
DT 19-DEC-2000 (first entry)  
XX  
A. caninum nematode-extracted anticoagulant protein AcanAP44.  
XX  
Nematode-extracted anticoagulant protein; AcanAP44; blood clotting;  
KW canine hookworm; thrombosis; vaccine.  
XX  
OS Ancylostoma caninum.  
XX  
PN US6087487-A.  
XX  
PD 11-JUL-2000.  
XX  
PF 12-FEB-1999; 99US-0249451.  
XX  
PR 17-OCT-1995; 95WO-US13231.  
PR 17-APR-1997; 97US-0809455.  
PR 18-OCT-1994; 94US-0326110.  
PR 05-JUN-1995; 95US-0461965.  
PR 05-JUN-1995; 95US-0465380.  
PR 05-JUN-1995; 95US-0486397.  
PR 05-JUN-1995; 95US-0486399.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Lauwereys MJ, Stanssens PEH, Jaspers LS, Gansemans YGJ, Moyle M;  
PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;  
XX  
WPI; 2000-531359/48.  
DR N-PSDB; AAA73379.  
XX  
PT New cDNA molecule encoding a protein having factor Xa inhibitory  
PT activity for preventing and treating blood clotting disorders,  
PT comprises nematode-extracted anticoagulant protein domains -  
XX  
PS Disclosure; Fig 16; 197pp; English.  
XX  
CC The present sequence is the Ancylostoma caninum nematode-extracted  
CC anticoagulant protein AcanAP44. Proteins of this kind have been shown to  
CC be effective at preventing blood clotting without causing excessive  
CC bleeding. The protein can be used in blood collection tubes to aid the  
CC isolation of plasma from the blood, to prevent thrombosis which may be  
CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
CC cancer and septic shock, and to produce antibodies. In the latter  
CC instance, the antibodies can be raised in order to detect infection by  
CC nematodes (the coding sequence can also be used for this) or as  
CC diagnostic tests. The protein can also be used as a vaccine against  
CC nematode parasites.  
XX  
SQ Sequence 86 AA;

Query Match 41.6%; Score 202; DB 21; Length 86;  
Best Local Similarity 47.4%; Pred. No. 8.3e-12;

Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;  
QY 5 OCGENEKYDSCGSKCKYDGVVEEDD--BEPNVPCLVRYCHQ--DCVCEEGFYRNK 60  
Db 5 KCGPGERLDCANKKPCPKIETSEEDDDVEETDVRCLVRVCERPLKICKDGYRNK 64  
QY 61 DDKCVSAEDCELDNMDPI 78  
Db 65 KGECVTDDVCQEDFMEFI 82  
Search completed: December 8, 2003, 09:16:38  
Job time : 114.545 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 39.0303 Seconds  
(without alignments)  
91.060 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMQCGENEKYDCSGKSC.....VSAEDCELDNMDFIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	2	US-08-465-380-59
2	486	100.0	84	2	US-08-486-397-59
3	486	100.0	84	2	US-08-486-399-59
4	486	100.0	84	2	US-08-461-965-59
5	486	100.0	84	2	US-08-634-641-59
6	486	100.0	84	3	US-09-249-471-59
7	486	100.0	84	3	US-09-249-472-59
8	486	100.0	84	3	US-09-249-451-59
9	486	100.0	84	3	US-08-809-455-59
10	486	100.0	84	3	US-09-249-461-59
11	486	100.0	84	3	US-09-249-448-59
12	486	100.0	84	4	US-09-249-473-59
13	486	100.0	91	2	US-08-465-380-128
14	486	100.0	91	2	US-08-480-478-50
15	486	100.0	91	2	US-08-486-397-128
16	486	100.0	91	2	US-08-486-399-128
17	486	100.0	91	2	US-08-461-965-128
18	486	100.0	91	2	US-08-326-110A-50
19	486	100.0	91	2	US-08-634-641-128
20	486	100.0	91	3	US-09-249-471-128
21	486	100.0	91	3	US-09-249-472-128
22	486	100.0	91	3	US-09-249-451-128
23	486	100.0	91	3	US-08-809-455-128
24	486	100.0	91	3	US-09-249-461-128
25	486	100.0	91	3	US-09-249-448-128
26	486	100.0	91	4	US-09-249-473-128
27	212	43.6	82	2	US-08-465-380-25

28	212	43.6	82	2	US-08-465-380-49	Sequence 49, Appl
29	212	43.6	82	2	US-08-480-478-54	Sequence 54, Appl
30	212	43.6	82	2	US-08-486-397-25	Sequence 25, Appl
31	212	43.6	82	2	US-08-486-399-49	Sequence 49, Appl
32	212	43.6	82	2	US-08-486-399-25	Sequence 25, Appl
33	212	43.6	82	2	US-08-486-399-49	Sequence 49, Appl
34	212	43.6	82	2	US-08-461-965-25	Sequence 25, Appl
35	212	43.6	82	2	US-08-461-965-49	Sequence 49, Appl
36	212	43.6	82	2	US-08-326-110A-54	Sequence 54, Appl
37	212	43.6	82	2	US-08-634-641-25	Sequence 25, Appl
38	212	43.6	82	2	US-08-634-641-49	Sequence 49, Appl
39	212	43.6	82	3	US-09-249-471-25	Sequence 25, Appl
40	212	43.6	82	3	US-09-249-471-49	Sequence 49, Appl
41	212	43.6	82	3	US-09-249-472-25	Sequence 25, Appl
42	212	43.6	82	3	US-09-249-472-49	Sequence 49, Appl
43	212	43.6	82	3	US-09-249-451-25	Sequence 25, Appl
44	212	43.6	82	3	US-09-249-451-49	Sequence 49, Appl
45	212	43.6	82	3	US-08-809-455-25	Sequence 25, Appl

## ALIGNMENTS

### RESULT 1

US-08-465-380-59  
; Sequence 59, Application US/08465380  
; Patent No 5963894  
; GENERAL INFORMATION:  
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
; APPLICANT: Yannick G.J. Ganssemaans, Matthew Moyle,  
; APPLICANT: Peter W. Bergum  
; TITLE OF INVENTION: NENATODE-EXTRACTED ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,380  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 213/268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum



US-08-461-965-59  
; Sequence 59, Application US/08461965  
; Patent No. 5872098  
; GENERAL INFORMATION:  
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
; APPLICANT: Peter W. Bergum  
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,965  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 210/243  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
; US-08-461-965-59

Query Match 100.0%; Score 486; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KATMOGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK-60  
Db 1 KATMOGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 5  
US-08-634-641-59  
; Sequence 59, Application US/08634641  
; Patent No. 5955294  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George P. Vlasuk  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Mensens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef

; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Gansemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew W.  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/634,641  
; FILING DATE: April 19, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 219/136  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
; US-08-634-641-59

Query Match 100.0%; Score 486; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMOGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK 60  
Db 1 KATMOGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK 60  
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 6  
US-09-249-471-59  
; Sequence 59, Application US/09249471  
; Patent No. 6040441

GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Larocche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Gansemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Berghum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,471  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum

US-09-249-471-59  
Query Match 100.0%; Score 486; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KATMCQGNENKDYSCGSKCKYKDYGVVEEDDEPNVPCLVRVCHODCVCEGFFYRNK 60  
DB 1 KATMCQGNENKDYSCGSKCKYKDYGVVEEDDEPNVPCLVRVCHODCVCEGFFYRNK 60  
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84

Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
RESULT 7  
US-09-249-472-59  
Sequence 59, Application US/09249472  
Patent No. 6046318  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Larocche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Gansemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Berghum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,472  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
US-09-249-472-59  
Query Match 100.0%; Score 486; DB 3; Length 84;

Best Local Similarity 100.0%; Pred. No. 4e-43; Indels 0; Gaps 0;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60  
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60

QY 61 DDKVSAEDCELDNMDFIYPGTRN 84  
Db 61 DDKVSAEDCELDNMDFIYPGTRN 84

RESULT 8  
US-09-249-451-59  
; Sequence 59, Application US/09249451  
; Patent No. 6087487  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jeppers, Laurent Stephane  
; APPLICANT: Gansemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,451  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
US-09-249-451-59

Query Match 100.0%; Score 486; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4e-43; Indels 0; Gaps 0;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60  
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60

QY 61 DDKVSAEDCELDNMDFIYPGTRN 84  
Db 61 DDKVSAEDCELDNMDFIYPGTRN 84

RESULT 9  
US-08-809-455-59  
; Sequence 59, Application US/08809455  
; Patent No. 6090916  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jeppers, Laurent Stephane  
; APPLICANT: Gansemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-809-455-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 10
US-09-249-461-59
; Sequence 59, Application US/09249461
; Patent No. 6096877
; GENERAL INFORMATION:
; APPLICANT: Vlaeuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,461
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
```

```
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-461-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 11
US-09-249-448-59
; Sequence 59, Application US/09249448
; Patent No. 6121435
; GENERAL INFORMATION:
; APPLICANT: Vlaeuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,448
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
```

```

; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
;
US-09-249-448-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 60
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 12
US-09-249-473-59
; Sequence 59, Application US/09249473
; Patent No. 6534629
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansmans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0

```

```

; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,473
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
;
US-09-249-473-59

Query Match 100.0%; Score 486; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 60
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEEPNVPCLVRVCHQDCVCEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 13
US-08-465-380-128
; Sequence 128, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansmans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

```

```
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-465-380-128

Query Match 100.0%; Score 486; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.3e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 8 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 67

Qy 61 DDKCVSAEDCCLDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCCLDNMDFIYPGTRN 91

RESULT 14
US-08-480-478-50
; Sequence 50, Application US/08480478
; Patent No. 5864009
; GENERAL INFORMATION:
; APPLICANT: GEORGE P. VLASUK; PATRICK ERIC
; APPLICANT: HUGO STANSSENS; JORIS HILDA
; APPLICANT: LIEVEN MESSENS; MARC JOZEF
; APPLICANT: LAUWEREYS; YVES RENE LAROCHE;
; APPLICANT: LAURENT STEPHANE JESPEERS; and
; APPLICANT: YANNICK GEORGES JOZEF
; APPLICANT: GANSEMAN
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-
; TITLE OF INVENTION: CONGLUTIN PROTEIN
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,478

; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-465-380-128

Query Match 100.0%; Score 486; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.3e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 8 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 67

Qy 61 DDKCVSAEDCCLDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCCLDNMDFIYPGTRN 91

RESULT 15
US-08-486-397-128
; Sequence 128, Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
```



TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 128:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
US-08-486-397-128

Query Match 100.0%; Score 486; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 4.3e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMOCCGENEKYDSCGSKCKKCKYDGVVEEDDEEPNVFCLVRVCHQDCVCEEGFYRNK 60  
Db 8 KATMOCCGENEKYDSCGSKCKKCKYDGVVEEDDEEPNVFCLVRVCHQDCVCEEGFYRNK 67

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91

Search completed: December 8, 2003, 09:20:28  
Job time : 39.0303 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 6.22222 Seconds  
(without alignments)  
209.232 Million cell updates/sec

Title: US-09-498-556C-79  
Perfect score: 32  
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	11	US-09-498-272-70
2	30	93.8	7	11	US-09-498-272-79
3	30	93.8	78	11	US-09-498-272-47
4	30	93.8	82	11	US-09-498-272-25
5	30	93.8	82	11	US-09-498-272-49
6	30	93.8	83	11	US-09-498-272-52
7	30	93.8	84	11	US-09-498-272-50
8	30	93.8	84	11	US-09-498-272-51
9	30	93.8	84	11	US-09-498-272-59
10	30	93.8	89	11	US-09-498-272-48
11	30	93.8	91	11	US-09-498-272-128
12	30	93.8	102	11	US-09-498-272-27
13	30	93.8	108	11	US-09-498-272-24
14	30	93.8	113	15	US-10-156-761-11215
15	30	93.8	138	10	US-09-764-877-1934

16	30	93.8	161	11	US-09-498-272-65
17	30	93.8	162	11	US-09-498-272-63
18	30	93.8	162	11	US-09-498-272-64
19	30	93.8	171	11	US-09-498-272-62
20	30	93.8	260	10	US-09-738-626-3828
21	30	93.8	268	10	US-09-864-866-43
22	30	93.8	475	12	US-10-205-219-65
23	30	93.8	533	15	US-10-207-655-75
24	30	93.8	1564	10	US-09-801-368-244
25	27	84.4	5	11	US-09-498-272-72
26	27	84.4	7	11	US-09-498-272-81
27	27	84.4	79	9	US-09-864-761-45609
28	27	84.4	86	11	US-09-498-272-45
29	27	84.4	86	11	US-09-498-272-46
30	27	84.4	87	11	US-09-498-272-44
31	27	84.4	88	11	US-09-498-272-43
32	27	84.4	173	9	US-09-925-302-855
33	27	84.4	260	10	US-09-858-546-5
34	27	84.4	392	10	US-09-813-718-16
35	27	84.4	415	10	US-09-813-718-14
36	27	84.4	437	10	US-09-813-718-12
37	27	84.4	471	11	US-09-919-039-163
38	27	84.4	471	12	US-10-247-671-166
39	27	84.4	471	15	US-10-126-467B-2
40	27	84.4	475	9	US-09-925-302-558
41	27	84.4	484	10	US-09-813-718-10
42	27	84.4	530	10	US-09-858-546-2
43	27	84.4	530	14	US-10-108-605-171
44	27	84.4	2771	10	US-09-808-602-82
45	27	84.4	2771	11	US-09-800-198-70

## ALIGNMENTS

## RESULT 1

US-09-498-272-70  
; Sequence 70, Application US/09498272  
; Publication No. US20030113890A1

## GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Ganssems, Yannick Georges Jozef  
Moyle, Matthew  
Berghum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: Los Angeles

COUNTRY: U.S.A.

ZIP: 90071

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486.399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486.397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465.380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461.965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326.110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
SEQUENCE DESCRIPTION: SEQ ID NO: 70:

Query Match 93.8%; Score 30; DB 11; Length 5;  
Best Local Similarity 100.0%; Pred. No. 6.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 1 GFYRN 5

## RESULT 2

US-09-498-272-79  
Sequence 79, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansemans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER READABLE FORM:  
Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486.399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486.397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465.380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461.965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326.110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
FEATURE:  
OTHER INFORMATION: Xaa in locations 1 and 2  
is an amino acid, provided  
that at least one Xaa is Glu or  
Asp.  
SEQUENCE DESCRIPTION: SEQ ID NO: 79:

US-09-498-272-79

Query Match 93.8%; Score 30; DB 11; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 3 GFYRN 7

## RESULT 3

US-09-498-272-47  
Sequence 47, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansemans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
MEDIUM TYPE: 3.5" diskette, 1.44 Mb  
COMPUTER READABLE FORM:  
storage

COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 78 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 47:  
US-09-498-272-47

Query Match 93.8%; Score 30; DB 11; Length 78;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 54 GFYRN 58

## RESULT 4

US-09-498-272-25  
Sequence 25, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jaspers, Laurent Stephane  
Ganseman, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma ceylanicum  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-498-272-25

Query Match 93.8%; Score 30; DB 11; Length 82;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 51 GFYRN 55

## RESULT 5

US-09-498-272-49  
Sequence 49, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jaspers, Laurent Stephane  
Ganseman, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles

STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma ceylanicum  
SEQUENCE DESCRIPTION: SEQ ID NO: 49:  
US-09-498-272-49

Query Match 93.8%; Score 30; DB 11; Length 82;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 51 GFYRN 55

## RESULT 6

US-09-498-272-52  
Sequence 52, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Laureys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansmans, Yannick Georges Jozef  
Moyle, Matthew  
Berghum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma duodenale  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
US-09-498-272-52

Query Match 93.8%; Score 30; DB 11; Length 83;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 55 GFYRN 59

## RESULT 7

US-09-498-272-50  
Sequence 50, Application US/09498272  
Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Laureys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansmans, Yannick Georges Jozef  
Moyle, Matthew  
Berghum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: California  
CITY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:

Query Match 93.8%; Score 30; DB 11; Length 84;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 56 GFYRN 60

## RESULT 8

US-09-498-272-51  
Sequence 51, Application US/09498272  
Publication No. US20030113890A1

## GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansmans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: California  
CITY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 51:

US-09-498-272-51

Query Match 93.8%; Score 30; DB 11; Length 84;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 56 GFYRN 60

## RESULT 9

US-09-498-272-59  
Sequence 59, Application US/09498272  
Publication No. US20030113890A1

## GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane

; Gansemans, Yannick Georges Jozef  
; Moyle, Matthew  
; Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; INHIBITORS AND ANTICOAGULANT  
; PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/498,272  
; FILING DATE: 04-Feb-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 216/270  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
; ORGANISM: Ancylostoma caninum  
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:  
US-09-498-272-59

Query Match 93.8%; Score 30; DB 11; Length 84;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 55 GFYRN 59  
|||||

RESULT 10  
US-09-498-272-48  
; Sequence 48, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Scansens, Patrick Eric Hugo  
; Messens, Joris Hilda Lieven

; Lauwereys, Marc Jozef  
; Laroche, Yves Rene  
; Jespers, Laurent Stephane  
; Gansemans, Yannick Georges Jozef  
; Moyle, Matthew  
; Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; INHIBITORS AND ANTICOAGULANT  
; PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/498,272  
; FILING DATE: 04-Feb-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 216/270  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 89 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma ceylanicum  
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
US-09-498-272-48

Query Match 93.8%; Score 30; DB 11; Length 89;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 61 GFYRN 65  
|||||

RESULT 11  
US-09-498-272-128  
; Sequence 128, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:



APPLICANT: Vlasuk, George Phillip  
 Stanssens, Patrick Eric Hugo  
 Messens, Joris Hilda Lieven  
 Lauwereys, Marc Josef  
 Laroche, Yves Rene  
 Jespers, Laurent Stephane  
 Gansmans, Yannick Georges Jozef  
 Moyle, Matthew  
 Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 INHIBITORS AND ANTICOAGULANT  
 PROTEIN

NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" diskette, 1.44 Mb  
 storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION NUMBER: US/09/498,272  
 FILING DATE: 04-Feb-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 128:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 91 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Ancylostoma caninum  
 SEQUENCE DESCRIPTION: SEQ ID NO: 128:  
 US-09-498-272-128

Query Match 93.8%; Score 30; DB 11; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 62 GFYRN 66

RESULT 12  
 US-09-498-272-27

Sequence 27, Application US/09498272  
 Publication No. US20030113890A1  
 GENERAL INFORMATION:  
 APPLICANT: Vlasuk, George Phillip  
 Stanssens, Patrick Eric Hugo  
 Messens, Joris Hilda Lieven  
 Lauwereys, Marc Josef  
 Laroche, Yves Rene  
 Jespers, Laurent Stephane  
 Gansmans, Yannick Georges Jozef  
 Moyle, Matthew  
 Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 INHIBITORS AND ANTICOAGULANT  
 PROTEIN  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/498,272  
 FILING DATE: 04-Feb-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/485,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 27:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 102 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Ancylostoma duodenale  
 SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
 US-09-498-272-27

Query Match 93.8%; Score 30; DB 11; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 74 GFYRN 78

RESULT 13  
US-09-498-272-24  
; Sequence 24, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Stanssens, Patrick Eric Hugo  
; Messens, Joris Hilda Lieven  
; Lauwereys, Marc Josef  
; Laroche, Yves Rene  
; Jespers, Laurent Stephane  
; Gansemaus, Yannick Georges Jozef  
; Moyle, Matthew  
; Bergum, Peter W.  
; TITLE OF INVENTION: NENATODE-EXTRACTED SERINE PROTEASE  
; INHIBITORS AND ANTICOAGULANT  
; PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/498,272  
; FILING DATE: 04-Feb-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/POCKET NUMBER: 216/270  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 108 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma ceylanicum  
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-09-498-272-24

Query Match 93.8%; Score 30; DB 11; Length 108;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

Db 80 GFYRN 84  
|||||

RESULT 14  
US-10-156-761-11215  
; Sequence 11215, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HAITORI, WASHIURA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 11215  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-11215

Query Match 93.8%; Score 30; DB 15; Length 113;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 51 GFYRN 55

RESULT 15  
US-09-764-877-1934  
; Sequence 1934, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1934  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-877-1934

Query Match 93.8%; Score 30; DB 10; Length 138;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 62 GFYRN 66

Search completed: December 8, 2003, 09:34:03  
Job time : 6.22222 secs



```
Query Match      93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
        |||||
Db      1 GFYRN 5

RESULT 2
US-08-486-397-70
; Sequence 70, Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
US-08-486-397-70

Query Match      93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
        |||||
Db      1 GFYRN 5

RESULT 3
US-08-486-399-70
; Sequence 70, Application US/08486399
```

```
; Patent No. 5866543
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,399
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
US-08-486-399-70

Query Match      93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
        |||||
Db      1 GFYRN 5

RESULT 4
US-08-461-965-70
; Sequence 70, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
```

STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,965  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 10, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 210/243  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-08-461-965-70

Query Match 93.8%; Score 30; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 1 GFYRN 5

RESULT 5  
US-08-634-641-70  
Sequence 70, Application US/08634641  
Patent No. 5955294  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George P. Vlasuk  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Mensens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jaspers, Laurent Stephane  
APPLICANT: Gansmans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/634,641  
FILING DATE: April 19, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 219/136  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-08-634-641-70

Query Match 93.8%; Score 30; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 1 GFYRN 5

RESULT 6  
US-09-249-471-70  
Sequence 70, Application US/09249471  
Patent No. 6040441  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jaspers, Laurent Stephane  
APPLICANT: Gansmans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage

;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: IBM P.C. DOS 5.0  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/249,471  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/809,455  
;; FILING DATE: April 17, 1997  
;; APPLICATION NUMBER: PCT/US95/13231  
;; FILING DATE: October 17, 1995  
;; APPLICATION NUMBER: 08/486,399  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/486,397  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/465,380  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/326,110  
;; FILING DATE: October 18, 1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BIGGS, SUZANNE L.  
;; REGISTRATION NUMBER: 30,158  
;; REFERENCE/DOCKET NUMBER: 216/270  
;; TELEPHONE: (213) 489-1600  
;; TELEX: 67-3510  
;; INFORMATION FOR SEQ ID NO: 70:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal fragment  
US-09-249-471-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 7  
US-09-249-472-70  
; Sequence 70, Application US/09249472  
; Patent No. 6046318  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Ganssemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Berqum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; STATE: Los Angeles  
; COUNTRY: California  
; ZIP: U.S.A.  
; ZIP: 90071

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
;; MEDIUM TYPE: Storage  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: IBM P.C. DOS 5.0  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/249,472  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/809,455  
;; FILING DATE: April 17, 1997  
;; APPLICATION NUMBER: PCT/US95/13231  
;; FILING DATE: October 17, 1995  
;; APPLICATION NUMBER: 08/486,399  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/486,397  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/465,380  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/461,965  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/326,110  
;; FILING DATE: October 18, 1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BIGGS, SUZANNE L.  
;; REGISTRATION NUMBER: 30,158  
;; REFERENCE/DOCKET NUMBER: 216/270  
;; TELEPHONE: (213) 489-1600  
;; TELEX: (213) 955-0440  
;; TELEEX: 67-3510  
;; INFORMATION FOR SEQ ID NO: 70:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal fragment  
US-09-249-472-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 8  
US-09-249-451-70  
; Sequence 70, Application US/09249451  
; Patent No. 6087487  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Ganssemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Berqum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; CITY: Los Angeles

STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,451  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-451-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPYRN 7  
Db 1 GPYRN 5

RESULT 9  
US-08-809-455-70  
Sequence 70, Application US/08809455  
Patent No. 6090916  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,455  
FILING DATE: April 17, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-08-809-455-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPYRN 7  
Db 1 GPYRN 5

RESULT 10  
US-09-249-461-70  
Sequence 70, Application US/09249461  
Patent No. 6096877  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Lyon & Lyon  
;; STREET: 633 West Fifth Street  
;; STREET: Suite 4700  
;; CITY: Los Angeles  
;; STATE: California  
;; COUNTRY: U.S.A.  
;; ZIP: 90071  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
;; MEDIUM TYPE: storage  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: IBM P.C. DOS 5.0  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/249,461  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/809,455  
;; FILING DATE: April 17, 1997  
;; APPLICATION NUMBER: PCT/US95/13231  
;; FILING DATE: October 17, 1995  
;; APPLICATION NUMBER: 08/486,399  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/461,965  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/326,110  
;; FILING DATE: October 18, 1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BIGGS, SUZANNE L.  
;; REGISTRATION NUMBER: 30,158  
;; REFERENCE/DOCKET NUMBER: 216/270  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (213) 489-1600  
;; TELEFAX: (213) 955-0440  
;; TELEX: 67-3510  
;; INFORMATION FOR SEQ ID NO: 70:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal fragment  
US-09-249-461-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 11  
US-09-249-448-70  
; Sequence 70, Application US/09249448  
; Patent No. 6121435  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Ganssemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE  
; INHIBITORS AND ANTICOAGULANT

;; TITLE OF INVENTION: PROTEIN  
;; NUMBER OF SEQUENCES: 356  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Lyon & Lyon  
;; STREET: 633 West Fifth Street  
;; STREET: Suite 4700  
;; CITY: Los Angeles  
;; STATE: California  
;; COUNTRY: U.S.A.  
;; ZIP: 90071  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
;; MEDIUM TYPE: storage  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: IBM P.C. DOS 5.0  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/249,448  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/809,455  
;; FILING DATE: April 17, 1997  
;; APPLICATION NUMBER: PCT/US95/13231  
;; FILING DATE: October 17, 1995  
;; APPLICATION NUMBER: 08/486,399  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/461,965  
;; FILING DATE: June 5, 1995  
;; APPLICATION NUMBER: 08/326,110  
;; FILING DATE: October 18, 1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BIGGS, SUZANNE L.  
;; REGISTRATION NUMBER: 30,158  
;; REFERENCE/DOCKET NUMBER: 216/270  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (213) 489-1600  
;; TELEFAX: (213) 955-0440  
;; TELEX: 67-3510  
;; INFORMATION FOR SEQ ID NO: 70:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal fragment  
US-09-249-448-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 12  
US-09-249-473-70  
; Sequence 70, Application US/09249473  
; Patent No. 6534629  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Ganssemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew



APPLICANT: Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/249,473  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/809,455  
 FILING DATE: April 17, 1997  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 70:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: internal fragment  
 US-09-249-473-70

Query Match 93.8%; Score 30; DB 4; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
 Db 1 GFYRN 5

RESULT 13  
 US-08-465-380-79  
 Sequence 79, Application US/08465380  
 Patent No. 5863894  
 GENERAL INFORMATION:  
 APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
 APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
 APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
 APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
 APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
 TITLE OF INVENTION: PROTEIN  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/465,380  
 FILING DATE: June 5, 1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 213/268  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 79:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: internal fragment  
 FEATURE:  
 OTHER INFORMATION: Xaa in locations 1 and 2  
 OTHER INFORMATION: is an amino acid, provided  
 OTHER INFORMATION: that at least one Xaa is Glu or  
 OTHER INFORMATION: Asp.  
 US-08-465-380-79

Query Match 93.8%; Score 30; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
 Db 3 GFYRN 7

RESULT 14  
 US-08-486-397-79  
 Sequence 79, Application US/08486397  
 Patent No. 5866542  
 GENERAL INFORMATION:  
 APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
 APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
 APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
 APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
 APPLICANT: Peter W. Bergum  
 TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
 NUMBER OF SEQUENCES: 357  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles

```
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FRAGMENT TYPE: internal fragment
; FEATURE:
; OTHER INFORMATION: Xaa in locations 1 and 2
; OTHER INFORMATION: is an amino acid, provided
; OTHER INFORMATION: that at least one Xaa is Glu or
; OTHER INFORMATION: Asp.
; US-08-486-399-79

Query Match 93.8%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 3 GFYRN 7

Search completed: December 8, 2003, 09:20:29
Job time: 4.25253 secs

; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FRAGMENT TYPE: internal fragment
; FEATURE:
; OTHER INFORMATION: Xaa in locations 1 and 2
; OTHER INFORMATION: is an amino acid, provided
; OTHER INFORMATION: that at least one Xaa is Glu or
; OTHER INFORMATION: Asp.
; US-08-486-397-79

Query Match 93.8%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 3 GFYRN 7

RESULT 15
US-08-486-399-79
; Sequence 79, Application US/08486399
; Patent No. 5866543
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssens, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 9.54545 Seconds  
(without alignments)  
116.399 Million cell updates/sec

Title: US-09-498-556C-79  
Perfect score: 32  
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	17	AA1991721
2	30	93.8	5	20	AA1930440
3	30	93.8	5	21	AA1915325
4	30	93.8	7	17	AA191730
5	30	93.8	7	20	AA1930449
6	30	93.8	7	21	AA1915334
7	30	93.8	78	20	AA1930420
8	30	93.8	78	21	AA1915305
9	30	93.8	82	20	AA1930399

10	30	93.8	82	20	AA1930422	Mature nematode ex
11	30	93.8	82	21	AA1915293	A. ceylanicum nema
12	30	93.8	82	21	AA1915307	A. ceylanicum nema
13	30	93.8	83	20	AA1930425	Mature nematode ex
14	30	93.8	84	20	AA1930423	Mature nematode ex
15	30	93.8	84	20	AA1930424	Mature nematode ex
16	30	93.8	84	20	AA1930432	Mature nematode ex
17	30	93.8	84	21	AA1915317	A. caninum nematod
18	30	93.8	86	22	AA191998	Human digestive sy
19	30	93.8	89	20	AA1930421	Mature nematode ex
20	30	93.8	91	17	AA191701	AcanAP2. Ancylos
21	30	93.8	91	20	AA1930393	Nematode extracted
22	30	93.8	91	20	AA1930454	Nematode extracted
23	30	93.8	91	21	AA1915346	A. caninum nematod
24	30	93.8	97	17	AA191709	AcanAP31. Ancylos
25	30	93.8	97	20	AA1930407	Nematode extracted
26	30	93.8	97	22	AA1914986	Human nervous syst
27	30	93.8	102	20	AA1930401	Nematode extracted
28	30	93.8	102	21	AA1915295	A. duodenale nemat
29	30	93.8	108	20	AA1930398	Nematode extracted
30	30	93.8	108	21	AA1915292	A. ceylanicum nema
31	30	93.8	124	21	AA1970530	Maize plastid targ
32	30	93.8	138	22	AA1903987	Human musculoskele
33	30	93.8	138	24	AA1913281	Novel human muscul
34	30	93.8	161	20	AA1930438	Mature nematode ex
35	30	93.8	162	20	AA1930436	Mature nematode ex
36	30	93.8	162	20	AA1930437	Mature nematode ex
37	30	93.8	162	21	AA1915321	A. caninum nematod
38	30	93.8	162	21	AA1915322	A. caninum nematod
39	30	93.8	171	20	AA1930435	Mature nematode ex
40	30	93.8	180	17	AA191713	AduNAP7. Ancylost
41	30	93.8	180	20	AA1930388	Nematode extracted
42	30	93.8	181	17	AA191711	AcanAP45. Ancylos
43	30	93.8	181	17	AA191712	AcanAP47. Ancylos
44	30	93.8	181	20	AA1930409	Nematode extracted
45	30	93.8	181	20	AA1930410	Nematode extracted

ALIGNMENTS

RESULT 1  
AA191721  
ID AA191721 standard; Protein; 5 AA.  
XX AA191721;  
AC AA191721;  
XX  
DT 25-MAR-2003 (updated)  
DT 17-NOV-1996 (first entry)  
XX  
DE NAP subsequence.  
XX  
KW AcanAP; HpoNAP; NamNAP; AcanAP; AduNAP; anticoagulant;  
KW nematode-extracted anticoagulant protein; serine protease;  
KW nematode; thrombosis; parasitic worm.  
XX  
OS Synthetic.  
XX  
PN WO9612021-A2.  
XX  
PD 25-APR-1996.  
XX  
PF 17-OCT-1995; 95WO-US13231.  
XX  
PR 05-JUN-1995; 95US-0486399.  
PR 18-OCT-1994; 94US-0326110.  
PR 05-JUN-1995; 95US-0461965.  
PR 05-JUN-1995; 95US-0465380.  
PR 05-JUN-1995; 95US-0486397.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Vlaauk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;

PI Jaspers LS, Gansemans YGJ, Moyle M, Bergum PW;  
 XX WPT; 1996-222007/22.  
 XX  
 PT Proteins with anticoagulant and/or serine protease inhibitory  
 PT activity - isolated from nematodes and useful to inhibit blood  
 PT coagulation  
 XX  
 PS Claim 10; Page 144; 243pp; English.  
 XX  
 CC Proteins with anticoagulant and/or serine protease inhibitory  
 CC activity, isolated from nematodes, are useful to inhibit blood  
 CC coagulation. The proteins can be added to blood collection tubes  
 CC defining the collection of mammalian plasma. They are also useful  
 CC to prevent or inhibit thrombosis, and may be given alone or in  
 CC combination with other therapeutic or in vivo diagnostic agents.  
 CC The proteins can serve as immunogens to raise antibodies for use in  
 CC the diagnosis and identification of NAP concn. levels in biological  
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.  
 CC They can also be used as immunogens in prophylactic and therapeutic  
 CC vaccines against parasitic worm infection. The proteins may  
 CC double the clotting time of human plasma in prothrombin time assays  
 CC when present at 10-50 nMol, and double the clotting time of human  
 CC plasma in activated partial thrombin time assays when present  
 CC at 10-100 nMol.  
 CC The anticoagulant proteins are pref. derived from  
 CC *Ancylostoma caninum*, *A. ceylanicum*, *A. duodenale*, *Necator*  
 CC *americanus* or *Heligmosomoides polygyrus*.  
 CC The proteins pref. have 2 NAP domains and specifically inhibit  
 CC the catalytic activity of the factor VIIa/TF complex in the  
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,  
 CC do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase.  
 CC Proteins given in AAR91720-R91732 are preferred subsequences  
 CC of a generic NAP sequence.  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 5 AA;

Query Match 93.8%; Score 30; DB 17; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 1 GFYRN 5

RESULT 2  
 AAY30440  
 ID AAY30440 standard; Peptide; 5 AA.  
 XX  
 AC AAY30440;  
 XX  
 DT 15-NOV-1999 (first entry)  
 XX  
 DE Nematode extracted anticoagulant protein fragment.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Unidentified.  
 XX  
 FN US5955294-A.  
 XX  
 PD 21-SEP-1999.  
 XX  
 PF 19-APR-1996; 96US-0634641.  
 XX  
 PR 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 XX Bergum PW, Gansemans YGJ, Jaspers LS, Laroche YR;  
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX  
 DR WPI; 1999-539569/45.  
 XX  
 PT Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains  
 XX  
 PS Disclosure; Column 151; 197pp; English.  
 XX  
 CC The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP) fragment. The protein has activity as an anticoagulant  
 CC and/or serine protease inhibitor. The protein contains at least one  
 CC NAP domain which has selective inhibitory activity for factor VIIa/TF.  
 CC The specification describes a method for screening an isolated protein  
 CC at least one domain for factor VIIa/TF selective inhibitory activity.  
 CC The method comprises determining the time to clotting effected by a  
 CC concentration of the isolated protein in an ex vivo prothrombin time  
 CC (PT) assay and an ex vivo activated partial thromboplastin time (aPTT)  
 CC assay; calculating prolongation of clotting effected by the isolated  
 CC protein in each of the PT and aPTT assay, with respect to a baseline  
 CC clotting value for each assay, where prolongation of clotting is  
 CC calculated as fold elevation of clotting time relative to a baseline  
 CC clotting value, where a doubling of clotting time is deemed a two-fold  
 CC elevation; and calculating a PT to aPTT prolongation ratio, where a  
 CC ratio at least one is indicative of factor VIIa/TF inhibitory activity.  
 CC The method is useful for determining if a protein has factor VIIa/TF  
 CC inhibitory activity.  
 XX  
 SQ Sequence 5 AA;

Query Match 93.8%; Score 30; DB 20; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 1 GFYRN 5

RESULT 3  
 AAB15325  
 ID AAB15325 standard; Peptide; 5 AA.  
 XX  
 AC AAB15325;  
 XX  
 DT 19-DEC-2000 (first entry)  
 XX  
 DE NAP domain fragment #2.  
 XX  
 KW Nematode-extracted anticoagulant protein; NAP domain; blood clotting;  
 KW canine hookworm; thrombosis; vaccine.  
 XX  
 OS Unidentified.  
 XX  
 FN US6087487-A.  
 XX  
 PD 11-JUL-2000.  
 XX  
 PF 12-FEB-1999; 99US-0249451.  
 XX  
 PR 17-OCT-1995; 95WO-US13231.  
 PR 17-APR-1997; 97US-0809455.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Ganssemans YGJ, Moyle M;  
 PI Bergum FW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.  
 DR  
 XX  
 XX New cDNA molecule encoding a protein having factor Xa inhibitory  
 PT activity for preventing and treating blood clotting disorders,  
 PT comprises nematode-extracted anticoagulant protein domains -  
 XX  
 XX Claim 3; Column 285; 197pp; English.  
 PS  
 XX The present sequence is a fragment of the NAP domain (see AAB15347),  
 CC which is found in all nematode-extracted anticoagulant proteins (NAPs).  
 CC Proteins of this kind have been shown to be effective at preventing  
 CC blood clotting without causing excessive bleeding. They can be used in  
 CC blood collection tubes to aid the isolation of plasma from the blood, to  
 CC prevent thrombosis which may be linked to the rupture of an  
 CC atherosclerotic plaque, acute myocardial infarction, angina,  
 CC thrombolytic therapy, percutaneous transluminal coronary angioplasty,  
 CC disseminated intravascular coagulopathy, infection, cancer and septic  
 CC shock, and to produce antibodies. In the latter instance, the antibodies  
 CC can be raised in order to detect infection by nematodes (the coding  
 CC sequences can also be used for this) or as diagnostic tests. The proteins  
 CC can also be used as vaccines against nematode parasites.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 93.8%; Score 30; DB 21; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 1 GFYRN 5  
 |||||  
 |||||  
 RESULT 4  
 AAR91730  
 ID AAR91730 standard; Protein; 7 AA.  
 XX  
 AC AAR91730;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 17-NOV-1996 (first entry)  
 DE  
 XX NAP subsequence.  
 XX  
 XX AcANAP; HponAP; NamNAP; AceNAP; AduNAP; anticoagulant;  
 KW nematode-extracted anticoagulant protein; serine protease;  
 KW nematode; thrombosis; parasitic worm.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label= Glu, Asp, OTHER  
 FT /note= "at least one of residue 1 or residue 2  
 FT is Glu or Asp"  
 FT  
 FT Misc-difference 2  
 FT /label= Glu, Asp, OTHER  
 FT /note= "at least one of residue 1 or residue 2  
 FT is Glu or Asp"  
 FT  
 XX WO9612021-A2.  
 PN  
 XX 25-APR-1996.  
 PD  
 XX 17-OCT-1995; 95WO-US13231.  
 PF

XX 05-JUN-1995; 95US-0486399.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 XX Vlasuk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;  
 PI Jespers LS, Ganssemans YGJ, Moyle M, Bergum FW;  
 XX WPI; 1996-222007/22.  
 DR  
 XX Proteins with anticoagulant and/or serine protease inhibitory  
 PT activity - isolated from nematodes and useful to inhibit blood  
 PT coagulation  
 XX  
 XX Claim 26; Page 147; 243pp; English.  
 PS  
 XX Proteins with anticoagulant and/or serine protease inhibitory  
 CC activity, isolated from nematodes, are useful to inhibit blood  
 CC coagulation. The proteins can be added to blood collection tubes  
 CC defining the collection of mammalian plasma. They are also useful  
 CC to prevent or inhibit thrombosis, and may be given alone or in  
 CC combination with other therapeutic or in vivo diagnostic agents.  
 CC The proteins can serve as immunogens to raise antibodies for use in  
 CC the diagnosis and identification of NAP concn. levels in biological  
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.  
 CC They can also be used as immunogens in prophylactic and therapeutic  
 CC vaccines against parasitic worm infection. The proteins may  
 CC double the clotting time of human plasma in prothrombin time assays  
 CC when present at 10-50 nMol, and double the clotting time of human  
 CC plasma in activated partial thrombin time assays when present  
 CC at 10-100 nMol.  
 CC The anticoagulant proteins are pref. derived from  
 CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator  
 CC americanus or Heligmosomoides polygyrus.  
 CC The proteins pref. have 2 NAP domains and specifically inhibit  
 CC the catalytic activity of the factor VIIa/TF complex in the  
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,  
 CC do not specifically inhibit the activation of factor VIII in the  
 CC absence of TF and do not specifically inhibit prothrombinase.  
 CC Proteins given in AAR91720-R91732 are preferred subsequences  
 CC of a generic NAP sequence.  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 7 AA;  
 Query Match 93.8%; Score 30; DB 17; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 3 GFYRN 7  
 |||||  
 |||||  
 RESULT 5  
 AAY30449  
 ID AAY30449 standard; Peptide; 7 AA.  
 XX  
 AC AAY30449;  
 XX  
 XX 15-NOV-1999 (first entry)  
 DT  
 XX Nematode extracted anticoagulant protein fragment.  
 DE  
 XX Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 KW  
 XX Unidentified.  
 OS  
 XX

```

PN US5955294-A.
XX
PD 21-SEP-1999.
XX
XX 19-APR-1996; 96US-0634641.
XX
XX 19-APR-1996; 96US-0634641.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
PA (CORV-) CORVAS INT INC.
XX
XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlasuk GP;
XX
XX WPI; 1999-539569/45.
DR
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains
XX
XX Disclosure; Column 155; 197pp; English.
XX
CC The present sequence represents a nematode extracted anticoagulant
CC protein (NAP) fragment. The protein has activity as an anticoagulant
CC and/or serine protease inhibitor. The protein contains at least one
CC NAP domain which has selective inhibitory activity for factor VIIa/TF.
CC The specification describes a method for screening an isolated protein
CC at least one domain for factor VIIa/TF selective inhibitory activity.
CC The method comprises determining the time to clotting effected by a
CC concentration of the isolated protein in an ex vivo prothrombin time
CC (PT) assay and an ex vivo activated partial thromboplastin time (aPTT)
CC assay; calculating prolongation of clotting effected by the isolated
CC protein in each of the PT and aPTT assay, with respect to a baseline
CC clotting value for each assay, where prolongation of clotting is
CC calculated as fold elevation of clotting time relative to a baseline
CC clotting value, and calculating a PT to aPTT prolongation ratio, where a
CC elevation; and calculating a PT to aPTT prolongation ratio, where a
CC ratio at least one is indicative of factor VIIa/TF inhibitory activity.
CC The method is useful for determining if a protein has factor VIIa/TF
CC inhibitory activity.
XX
SQ Sequence 7 AA;
Query Match 93.8%; Score 30; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 3 GFYRN 7
|||
|
RESULT 6
AAB15334
ID AAB15334 standard; Peptide; 7 AA.
XX
XX AAB15334;
AC AAB15334;
XX
XX 19-DEC-2000 (first entry)
DT
XX NAP domain fragment #11.
DE
XX Nematode-extracted anticoagulant protein; NAP domain; blood clotting;
KW canine hookworm; thrombosis; vaccine.
XX
XX Unidentified.
OS
XX Key Location/Qualifiers
PH

```

---

```

FT Misc-difference 1..2
FT /label= OTHER
FT /note= "one of these two residues must be either
FT Glu or Asp"
XX
XX US6087487-A.
XX
PD 11-JUL-2000.
XX
XX 12-FEB-1999; 99US-0249451.
XX
XX 17-OCT-1995; 95WO-US13231.
PR 17-APR-1997; 97US-0809455.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
XX
XX (CORV-) CORVAS INT INC.
PA
XX
XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;
XX
XX WPI; 2000-531359/48.
DR
XX
XX New cDNA molecule encoding a protein having factor Xa inhibitory
PT activity for preventing and treating blood clotting disorders,
PT comprises nematode-extracted anticoagulant protein domains -
XX
XX Claim 4; Column 286; 197pp; English.
XX
CC The present sequence is a fragment of the NAP domain (see AAB15347),
CC which is found in all nematode-extracted anticoagulant proteins (NAPs).
CC Proteins of this kind have been shown to be effective at preventing
CC blood clotting without causing excessive bleeding. They can be used in
CC blood collection tubes to aid the isolation of plasma from the blood, to
CC prevent thrombosis which may be linked to the rupture of an
CC atherosclerotic plaque, acute myocardial infarction, angina,
CC thrombolytic therapy, percutaneous transluminal coronary angioplasty,
CC disseminated intravascular coagulopathy, infection, cancer and septic
CC shock, and to produce antibodies. In the latter instance, the antibodies
CC can be raised in order to detect infection by nematodes (the coding
CC sequences can also be used for this) or as diagnostic tests. The proteins
CC can also be used as vaccines against nematode parasites.
XX
SQ Sequence 7 AA;
Query Match 93.8%; Score 30; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 3 GFYRN 7
|||
|
RESULT 7
AAY30420
ID AAY30420 standard; Protein; 78 AA.
XX
XX AAY30420;
AC AAY30420;
XX
XX 15-NOV-1999 (first entry)
DT
XX Mature nematode extracted anticoagulant protein AcanAP31.
DE
XX Nematode extracted anticoagulant protein; NAP, anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
XX
XX Ancylostoma caninum.
OS
XX US5955294-A.
PN

```

XX PD 21-SEP-1999.  
 XX PF 19-APR-1996; 96US-0634641.  
 XX PR 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX (CORV-) CORVAS INT INC.  
 XX PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX WPI; 1999-539569/45.  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains  
 PT Disclosure; Columns 133-134; 197pp; English.  
 XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of  
 CC the PT and APTT assay, with respect to a baseline clotting value for  
 CC each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where  
 CC a doubling of clotting time is deemed a two-fold elevation; and  
 CC calculating a PT to APTT prolongation ratio, where a ratio at least  
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
 CC useful for determining if a protein has factor VIIa/TF inhibitory  
 CC activity.  
 XX SQ Sequence 78 AA;  
 Query Match 93.8%; Score 30; DB 20; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 54 GFYRN 58  
 RESULT 8  
 AAB15305  
 ID AAB15305 standard; Protein; 78 AA.  
 XX AC AAB15305;  
 XX 19-DEC-2000 (first entry)  
 DT A. caninum nematode-extracted anticoagulant protein AcanAP31,42,46.  
 DE Nematode-extracted anticoagulant protein; AcanAP31; AcanAP42; AcanAP46;  
 KW canine hookworm; blood clotting; thrombosis; vaccine.  
 XX Ancylostoma caninum.  
 OS US6087487-A.  
 XX PN

PD 11-JUL-2000.  
 XX 12-FEB-1999; 99US-0249451.  
 XX PR 17-OCT-1995; 95WO-US13231.  
 PR 17-APR-1997; 97US-0809455.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 XX (CORV-) CORVAS INT INC.  
 XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.  
 DR N-PSDB; AAA73378.  
 XX New cDNA molecule encoding a protein having factor Xa inhibitory  
 PT activity for preventing and treating blood clotting disorders,  
 PT comprises nematode-extracted anticoagulant protein domains -  
 XX Disclosure; Fig 16; 197pp; English.  
 XX The present sequence comprises the Ancylostoma caninum  
 CC nematode-extracted anticoagulant proteins AcanAP31, AcanAP42 and  
 CC AcanAP46. Proteins of this kind have been shown to be effective at  
 CC preventing blood clotting without causing excessive bleeding. The  
 CC proteins can be used in blood collection tubes to aid the isolation of  
 CC plasma from the blood, to prevent thrombosis which may be linked to the  
 CC rupture of an atherosclerotic plaque, acute myocardial infarction,  
 CC angina, thrombolytic therapy, percutaneous transluminal coronary  
 CC angioplasty, disseminated intravascular coagulopathy, infection, cancer  
 CC and septic shock, and to produce antibodies. In the latter instance, the  
 CC antibodies can be raised in order to detect infection by nematodes (the  
 CC coding sequence can also be used for this) or as diagnostic tests. The  
 CC proteins can also be used as a vaccine against nematode parasites.  
 XX SQ Sequence 78 AA;  
 Query Match 93.8%; Score 30; DB 21; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 54 GFYRN 58  
 RESULT 9  
 AAY30399  
 ID AAY30399 standard; Protein; 82 AA.  
 XX AC AAY30399;  
 XX 15-NOV-1999 (first entry)  
 DT Nematode extracted anticoagulant protein AcanAP4d2.  
 DE Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX Ancylostoma ceylanicum.  
 OS US5955294-A.  
 XX 21-SEP-1999.  
 PD 19-APR-1996; 96US-0634641.  
 XX 19-APR-1996; 96US-0634641.

```
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlasuk GP;
XX
XX WPI; 1999-539569/45.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains
PT
XX
XX Disclosure; Column 111-112; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (APTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of
CC the PT and APTT assay, with respect to a baseline clotting value for
CC each assay, where prolongation of clotting is calculated as fold
CC elevation of clotting time relative to a baseline clotting value, where
CC a doubling of clotting time is deemed a two-fold elevation; and
CC calculating a PT to APTT prolongation ratio, where a ratio at least
CC one is indicative of factor VIIa/TF inhibitory activity. The method is
CC useful for determining if a protein has factor VIIa/TF inhibitory
CC activity.
XX
XX SQ Sequence 82 AA;
Query Match 93.8%; Score 30; DB 20; Length 82;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 51 GFYRN 55
|||||
RESULT 10
AAAY30422
ID AAAY30422 standard; Protein; 82 AA.
XX
XX AAAY30422;
XX
XX 15-NOV-1999 (first entry)
XX
XX Mature nematode extracted anticoagulant protein AceNAP4d2.
DE
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
KW
XX
XX Ancylostoma ceylanicum.
OS
XX
XX US5955294-A.
PN
XX
XX 21-SEP-1999.
PD
XX
XX 19-APR-1996; 96US-0634641.
XX
XX 19-APR-1996; 96US-0634641.
PR
XX 18-OCT-1994; 94US-0326110.
PR
```

---

```
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlasuk GP;
XX
XX WPI; 1999-539569/45.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains
PT
XX
XX Disclosure; Columns 135-136; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (APTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of
CC the PT and APTT assay, with respect to a baseline clotting value for
CC each assay, where prolongation of clotting is calculated as fold
CC elevation of clotting time relative to a baseline clotting value, where
CC a doubling of clotting time is deemed a two-fold elevation; and
CC calculating a PT to APTT prolongation ratio, where a ratio at least
CC one is indicative of factor VIIa/TF inhibitory activity. The method is
CC useful for determining if a protein has factor VIIa/TF inhibitory
CC activity.
XX
XX SQ Sequence 82 AA;
Query Match 93.8%; Score 30; DB 20; Length 82;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 51 GFYRN 55
|||||
RESULT 11
AAB15293
ID AAB15293 standard; Protein; 82 AA.
XX
XX AAB15293;
XX
XX 19-DEC-2000 (first entry)
XX
XX A. ceylanicum nematode-extracted anticoagulant protein AceNAP4d2 #1.
DE
XX Nematode-extracted anticoagulant protein; AceNAP4d2; blood clotting;
KW canine hookworm; thrombosis; vaccine.
KW
XX
XX Ancylostoma ceylanicum.
OS
XX
XX US6087487-A.
PN
XX
XX 11-JUL-2000.
PD
XX
XX 12-FEB-1999; 99US-0249451.
XX
XX 17-OCT-1995; 95WO-US13231.
PR
XX 17-APR-1997; 97US-0809455.
PR
XX 18-OCT-1994; 94US-0326110.
PR
```



PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.  
 DR  
 XX  
 XX  
 PT New cDNA molecule encoding a protein having factor Xa inhibitory  
 PT activity for preventing and treating blood clotting disorders,  
 PT comprises nematode-extracted anticoagulant protein domains -  
 XX  
 XX Disclosure; Fig 11; 197pp; English.  
 XX  
 CC The present sequence is the Ancylostoma ceylanicum nematode-extracted  
 CC anticoagulant protein AcenAP4d2. Proteins of this kind have been shown to  
 CC be effective at preventing blood clotting without causing excessive  
 CC bleeding. The protein can be used in blood collection tubes to aid the  
 CC isolation of plasma from the blood, to prevent thrombosis which may be  
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 CC cancer and septic shock, and to produce antibodies. In the latter  
 CC instance, the antibodies can be raised in order to detect infection by  
 CC nematodes or as diagnostic tests. The protein can also be used as a  
 CC vaccine against nematode parasites.  
 XX  
 SQ Sequence 82 AA;  
 Query Match 93.8%; Score 30; DB 21; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 51 GFYRN 55  
 |||||  
 RESULT 12  
 AABI5307  
 ID AABI5307 standard; Protein; 82 AA.  
 AC AABI5307;  
 XX  
 DT 19-DEC-2000 (first entry)  
 DE  
 DE A. ceylanicum nematode-extracted anticoagulant protein AcenAP4d2 #2.  
 XX  
 KW Nematode-extracted anticoagulant protein; AcenAP4d2; blood clotting;  
 KW canine hookworm; thrombosis; vaccine.  
 XX  
 OS Ancylostoma ceylanicum.  
 XX  
 PN US6087487-A.  
 XX  
 XX  
 PD 11-JUL-2000.  
 XX  
 XX 12-FEB-1999; 99US-0249451.  
 XX  
 PR 17-OCT-1995; 95WO-US13231.  
 PR 17-APR-1997; 97US-0809455.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.  
 DR  
 XX  
 XX  
 PT New cDNA molecule encoding a protein having factor Xa inhibitory  
 PT activity for preventing and treating blood clotting disorders,  
 PT comprises nematode-extracted anticoagulant protein domains -  
 XX  
 XX Disclosure; Fig 11; 197pp; English.  
 XX  
 CC The present sequence is the Ancylostoma ceylanicum nematode-extracted  
 CC anticoagulant protein AcenAP4d2. Proteins of this kind have been shown to  
 CC be effective at preventing blood clotting without causing excessive  
 CC bleeding. The protein can be used in blood collection tubes to aid the  
 CC isolation of plasma from the blood, to prevent thrombosis which may be  
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 CC cancer and septic shock, and to produce antibodies. In the latter  
 CC instance, the antibodies can be raised in order to detect infection by  
 CC nematodes or as diagnostic tests. The protein can also be used as a  
 CC vaccine against nematode parasites.  
 XX  
 SQ Sequence 82 AA;  
 Query Match 93.8%; Score 30; DB 21; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 51 GFYRN 55  
 |||||  
 RESULT 13  
 AAY30425  
 ID AAY30425 standard; Protein; 83 AA.  
 XX  
 AC AAY30425;  
 XX  
 DT 15-NOV-1999 (first entry)  
 DE  
 DE Mature nematode extracted anticoagulant protein AduNAP7d1.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Ancylostoma duodenale.  
 XX  
 PN US5955294-A.  
 XX  
 PD 21-SEP-1999.  
 XX  
 XX 19-APR-1996; 96US-0634641.  
 XX  
 PR 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX WPI; 1999-539569/45.  
 DR  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant

PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.  
 DR  
 XX  
 XX  
 PT New cDNA molecule encoding a protein having factor Xa inhibitory  
 PT activity for preventing and treating blood clotting disorders,  
 PT comprises nematode-extracted anticoagulant protein domains -  
 XX  
 XX Disclosure; Fig 16; 197pp; English.  
 XX  
 CC The present sequence is the Ancylostoma ceylanicum nematode-extracted  
 CC anticoagulant protein AcenAP4d2. Proteins of this kind have been shown to  
 CC be effective at preventing blood clotting without causing excessive  
 CC bleeding. The protein can be used in blood collection tubes to aid the  
 CC isolation of plasma from the blood, to prevent thrombosis which may be  
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 CC cancer and septic shock, and to produce antibodies. In the latter  
 CC instance, the antibodies can be raised in order to detect infection by  
 CC nematodes or as diagnostic tests. The protein can also be used as a  
 CC vaccine against nematode parasites.  
 XX  
 SQ Sequence 82 AA;  
 Query Match 93.8%; Score 30; DB 21; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 51 GFYRN 55  
 |||||  
 RESULT 13  
 AAY30425  
 ID AAY30425 standard; Protein; 83 AA.  
 XX  
 AC AAY30425;  
 XX  
 DT 15-NOV-1999 (first entry)  
 DE  
 DE Mature nematode extracted anticoagulant protein AduNAP7d1.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Ancylostoma duodenale.  
 XX  
 PN US5955294-A.  
 XX  
 PD 21-SEP-1999.  
 XX  
 XX 19-APR-1996; 96US-0634641.  
 XX  
 PR 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0461965.  
 PR 05-JUN-1995; 95US-0465380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486399.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX WPI; 1999-539569/45.  
 DR  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains  
XX  
PS Disclosure; Columns 137-138; 197pp; English.  
XX  
CC The present sequence represents a nematode extracted anticoagulant protein (NAP). The protein has activity as an anticoagulant and/or serine protease inhibitor. The protein contains at least one NAP domain which has selective inhibitory activity for factor VIIa/TF. The specification describes a method for screening an isolated protein at least one domain for factor VIIa/TF selective inhibitory activity. The method comprises determining the time to clotting effected by a concentration of the isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo activated partial thromboplastin time (aPTT) assay; calculating prolongation of clotting effected by the isolated protein in each of the PT and aPTT assay, with respect to a baseline clotting value for each assay, where prolongation of clotting is calculated as fold elevation of clotting time relative to a baseline clotting value, where a doubling of clotting time is deemed a two-fold elevation; and calculating a PT to aPTT prolongation ratio, where a ratio at least one is indicative of factor VIIa/TF inhibitory activity. The method is useful for determining if a protein has factor VIIa/TF inhibitory activity.

XX  
SQ Sequence 83 AA;  
Query Match 93.8%; Score 30; DB 20; Length 83;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 55 GFYRN 59  
|||||

RESULT 14  
AAY30423  
ID AAY30423 standard; Protein; 84 AA.  
XX  
AC AAY30423;  
XX  
DT 15-NOV-1999 (first entry)  
XX  
DE Mature nematode extracted anticoagulant protein AcaNAP45d1.  
XX  
KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
XX  
OS Ancylostoma caninum.  
XX  
PN US9595294-A.  
XX  
PD 21-SEP-1999.  
XX  
PF 19-APR-1996; 96US-0634641.  
XX  
PR 19-APR-1996; 96US-0634641.  
PR 18-OCT-1994; 94US-0326110.  
PR 05-JUN-1995; 95US-0461965.  
PR 05-JUN-1995; 95US-0465380.  
PR 05-JUN-1995; 95US-0486397.  
PR 05-JUN-1995; 95US-0486399.  
PR 17-OCT-1995; 95WO-US13231.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
PI Vlaeuk GP;  
XX  
XX WPI; 1999-539569/45.  
DR  
XX Screening an isolated protein for Nematode-extracted Anticoagulant Protein domains  
PT  
XX

XX  
PS Disclosure; Columns 135-136; 197pp; English.  
XX  
CC The present sequence represents a nematode extracted anticoagulant protein (NAP). The protein has activity as an anticoagulant and/or serine protease inhibitor. The protein contains at least one NAP domain which has selective inhibitory activity for factor VIIa/TF. The specification describes a method for screening an isolated protein at least one domain for factor VIIa/TF selective inhibitory activity. The method comprises determining the time to clotting effected by a concentration of the isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo activated partial thromboplastin time (aPTT) assay; calculating prolongation of clotting effected by the isolated protein in each of the PT and aPTT assay, with respect to a baseline clotting value for each assay, where prolongation of clotting is calculated as fold elevation of clotting time relative to a baseline clotting value, where a doubling of clotting time is deemed a two-fold elevation; and calculating a PT to aPTT prolongation ratio, where a ratio at least one is indicative of factor VIIa/TF inhibitory activity. The method is useful for determining if a protein has factor VIIa/TF inhibitory activity.

XX  
SQ Sequence 84 AA;  
Query Match 93.8%; Score 30; DB 20; Length 84;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 56 GFYRN 60  
|||||

RESULT 15  
AAY30424  
ID AAY30424 standard; Protein; 84 AA.  
XX  
AC AAY30424;  
XX  
DT 15-NOV-1999 (first entry)  
XX  
DE Mature nematode extracted anticoagulant protein AcaNAP47d1.  
XX  
KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
XX  
OS Ancylostoma caninum.  
XX  
PN US9595294-A.  
XX  
PD 21-SEP-1999.  
XX  
PF 19-APR-1996; 96US-0634641.  
XX  
PR 19-APR-1996; 96US-0634641.  
PR 18-OCT-1994; 94US-0326110.  
PR 05-JUN-1995; 95US-0461965.  
PR 05-JUN-1995; 95US-0465380.  
PR 05-JUN-1995; 95US-0486397.  
PR 05-JUN-1995; 95US-0486399.  
PR 17-OCT-1995; 95WO-US13231.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;  
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
PI Vlaeuk GP;  
XX  
XX WPI; 1999-539569/45.  
DR  
XX Screening an isolated protein for Nematode-extracted Anticoagulant Protein domains  
PT  
XX

PS Disclosure; Columns 135-138; 197pp; English.

XX  
CC The present sequence represents a nematode extracted anticoagulant  
CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
CC protease inhibitor. The protein contains at least one NAP domain which  
CC has selective inhibitory activity for factor VIIa/TF. The specification  
CC describes a method for screening an isolated protein at least one domain  
CC for factor VIIa/TF selective inhibitory activity. The method comprises  
CC determining the time to clotting effected by a concentration of the  
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
CC activated partial thromboplastin time (aPTT) assay; calculating  
CC prolongation of clotting effected by the isolated protein in each of  
CC the PT and aPTT assay, with respect to a baseline clotting value for  
CC each assay, where prolongation of clotting is calculated as fold  
CC elevation of clotting time relative to a baseline clotting value, where  
CC a doubling of clotting time is deemed a two-fold elevation; and  
CC calculating a PT to aPTT prolongation ratio, where a ratio at least  
CC one is indicative of factor VIIa/TF inhibitory activity. The method is  
CC useful for determining if a protein has factor VIIa/TF inhibitory  
CC activity.

XX  
SQ Sequence 84 AA;

Query Match 93.8%; Score 30; DB 20; Length 84;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||  
Db 56 GFYRN 60

Search completed: December 8, 2003, 09:16:39  
Job time : 10.5455 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.83838 Seconds  
(without alignments)  
179.063 Million cell updates/sec

Title: US-09-498-556C-79  
Perfect score: 32  
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	93.8	121	1 MTR3 MOUSE	Q92d0 mus musculus
2	30	93.8	175	1 HPRT_MYCGE	P47696 mycoplasma
3	30	93.8	175	1 HPRT_MYCPN	P75119 mycoplasma
4	30	93.8	245	1 END3 MYCTU	O69642 mycobacteri
5	30	93.8	279	1 UVEN MYCLU	P46303 micrococcus
6	30	93.8	361	1 SERC_XANAC	Q8p1y7 xanthomonas
7	30	93.8	361	1 SERC_XANCP	Q8p897 xanthomonas
8	30	93.8	361	1 SERC_YERPE	Q8zgb4 yersinia pe
9	30	93.8	362	1 SERC_XYLFA	Q9pb19 xyella fas
10	30	93.8	378	1 SERC_RALSO	Q8y0z0 ralstonia s
11	30	93.8	423	1 PUR2 RHIME	Q92r10 rhizobium m
12	30	93.8	481	1 SYW_MOUSE	P32921 mus musculu
13	30	93.8	533	1 TYRO_MOUSE	P11344 mus musculu
14	30	93.8	824	1 YOT5_CABEL	P34651 caenorhabdi
15	30	93.8	900	1 XPC_MOUSE	P51612 mus musculu
16	30	93.8	939	1 XPC_HUMAN	Q01831 homo sapien
17	30	93.8	1564	1 PDRA_YEAST	P51533 saccharomyc
18	27	84.4	59	1 RL32_BORBU	O51646 borrelia bu
19	27	84.4	201	1 YAFJ_LACLA	Q9c1b0 lactococcus
20	27	84.4	218	1 RAD_CHLTE	Q8k669 chlorobium
21	27	84.4	234	1 DEOD_BUCAI	P57606 buchnera ap
22	27	84.4	299	1 YF35_METJA	Q58930 methanococc
23	27	84.4	318	1 SOLR_CIOAB	P33746 clostridium
24	27	84.4	320	1 K6PF_BUCAP	Q8k9n0 buchnera ap
25	27	84.4	340	1 YD17_ARCFU	Q28952 archaeoglob
26	27	84.4	344	1 Y613_METJA	Q58030 methanococc
27	27	84.4	465	1 YCJX_ECOLI	P76046 escherichia
28	27	84.4	471	1 SYW_HUMAN	P23381 homo sapien
29	27	84.4	471	1 YAL5_SCHPO	Q09928 schizosach
30	27	84.4	475	1 SYW_BOVIN	P17248 bos taurus
31	27	84.4	475	1 SYW_RABIT	P23612 oryctolagus
32	27	84.4	492	1 ANKH_MOUSE	Q9jh22 mus musculu
33	27	84.4	492	1 ANKH_RAT	P58366 rattus norv

34 27 84.4 550 1 PHNL\_DESGI P12944 desulfovibr  
35 27 84.4 608 1 SVT\_PYRAE Q82wk4 pyrobaculum  
36 27 84.4 667 1 TSI1\_GIALA Q03185 giardia lam  
37 27 84.4 713 1 TSN4\_GIALA P21849 giardia lam  
38 26 81.2 91 1 V179\_FOWPV Q9J554 fowlpox vir  
39 26 81.2 194 1 RK9\_PEA P11894 pisum sativ  
40 26 81.2 197 1 RK9\_ARATH P25864 arabidopsis  
41 26 81.2 218 1 YAFI\_LACLA Q9c1b5 lactococcus  
42 26 81.2 219 1 END3\_SYNV3 P71715 synecocyst  
43 26 81.2 245 1 END3\_MYCLE Q9cb92 mycobacteri  
44 26 81.2 281 1 YQAC\_LACLA Q9c1b9 lactococcus  
45 26 81.2 282 1 YF57\_STRPN Q97pp6 streptococc

## ALIGNMENTS

RESULT 1  
MTR3\_MOUSE  
ID MTR3\_MOUSE STANDARD; PRT; 121 AA.  
AC Q92ZD0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Myotubularin-related protein 3 (Fragment).  
GN MTR3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98409499; PubMed=9736772;  
RA Laporte J., Buj-Bello A., Tentler D., Kretz C., Dahl N.,  
RA Mandel J.-L.;  
RT "Characterization of the myotubularin dual specificity phosphatase  
RT gene family from yeast to human."  
RL Hum. Mol. Genet. 7:1703-1712(1998).  
CC -!- FUNCTION: NOT KNOWN. COULD BE A TYROSINE-PHOSPHATASE.  
CC -!- SIMILARITY: BELONGS TO THE MYOTUBULARIN-LIKE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF073881; AAC80003.1; -;  
DR MGD; MGI:1921552; Mtr3.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; PARTIAL.  
DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; PARTIAL.  
FT Hydrolase.  
KW NON\_TER  
SQ SEQUENCE 121 AA; 14292 MW; 0ECA07469DABB98 CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred.No. 8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
Db 105 GFYRN 109  
RESULT 2  
HPRT\_MYCGE STANDARD; PRT; 175 AA.  
ID HPRT\_MYCGE  
AC P47696;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPRT)
DE (HGPRTase).
GN HPT OR MG458.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Dougherty T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Utterback B.A., Boff K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 9-98 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Boff K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -|- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- CATALYTIC ACTIVITY: GMP + diphosphate = guanine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- PATHWAY: Purine salvage.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39727; AAC72478.1; -.
DR EMBL; U02193; AAD12479.1; -.
DR PIR; F64250; F64250.
DR HSP; P51900; LHGX.
DR TIGR; MG458; -.
DR InterPro; IPR005904; Hxn_phospho_trans.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR InterPro; IPR000836; Prtransferase.
DR Pfam; PF00156; Pribo syltran; 1.
DR TIGRfam; TIGR01203; HGPRTase; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage; Magnesium;
FT METAL 96 96 MAGNESIUM (BY SIMILARITY).
FT METAL 97 97 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 175 AA; 19759 MW; 9647332FF3E18E0 CRC64;
Query Match 93.8%; Score 30; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 GFYRN 7
Db 159 GFYRN 163
|||||
RESULT 3
HPRT_MYCPN STANDARD; PRT; 175 AA.
ID HPRT_MYCPN STANDARD; PRT; 175 AA.
```

```
AC P75119;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPRT)
DE (HGPRTase).
GN HPT OR MPN672 OR MP170.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plegens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -|- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- CATALYTIC ACTIVITY: GMP + diphosphate = guanine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- PATHWAY: Purine salvage.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A5000018; AAB95818.1; -.
DR PIR; S73496; S73496.
DR HSP; P00492; IHMP.
DR InterPro; IPR005904; Hxn_phospho_trans.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR InterPro; IPR000836; Prtransferase.
DR Pfam; PF00156; Pribo syltran; 1.
DR TIGRfam; TIGR01203; HGPRTase; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage; Magnesium;
FT METAL 96 96 MAGNESIUM (BY SIMILARITY).
FT METAL 97 97 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 175 AA; 19620 MW; 1E5A9FAA01D69854 CRC64;
Query Match 93.8%; Score 30; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 GFYRN 7
Db 159 GFYRN 163
|||||
RESULT 4
END3_MYCTU STANDARD; PRT; 245 AA.
ID END3_MYCTU STANDARD; PRT; 245 AA.
AC O69642;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endonuclease III (EC 4.2.99.18) (DNA-(apurinic or apyrimidinic site)
DE lyase).
GN NTH OR RV3674C OR MT3775 OR MTU025.022C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
```

OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Blehah W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: HAS BOTH AN APURINIC AND/OR APYRIMIDINIC ENDONUCLEASE  
CC ACTIVITY AND A DNA N-GLYCOSYLASE ACTIVITY. INCISES DAMAGED DNA AT  
CC CYTOSINES, THYMINES AND GUANINES. ACTS ON A DAMAGED STRAND, 5'  
CC FROM THE DAMAGED SITE (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or  
CC apyrimidinic site in DNA is broken by a beta-elimination reaction,  
CC leaving a 3'-terminal unsaturated sugar and a product with a  
CC terminal 5'-phosphate.  
CC -1- COFACTOR: BINDS A 4FE-4S CLUSTER WHICH IS NOT IMPORTANT FOR THE  
CC CATALYTIC ACTIVITY, BUT WHICH IS PROBABLY INVOLVED IN THE PROPER  
CC POSITIONING OF THE ENZYME ALONG THE DNA STRAND (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NTH/MUTY FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AL022121; CRA17996.1; ALT INIT.  
CC EMBL; AE007175; AAK48142.1; ALT\_INIT.  
CC HSSP; P20625; 2ABK.  
CC TIGR; MT3775; --  
CC Tuberculaist; RV3674c; --  
CC InterPro; IPR003265; Endo\_3c.  
CC InterPro; IPR004035; EndoIII\_FCL.  
CC InterPro; IPR004036; EndoIII\_HhH.  
CC InterPro; IPR003651; FeS bind.  
CC InterPro; IPR000445; HhH.  
CC InterPro; IPR003583; HhH\_1.  
CC InterPro; IPR005759; Nth.  
CC Pfam; PF00730; HhH-GPD; 1.  
CC Pfam; PF00633; HhH; 1.  
CC SMART; SM00478; ENDO3c; 1.  
CC SMART; SM00525; FES; 1.  
CC SMART; SM00278; HhH1; 1.  
CC TIGRfams; TIGR01083; nth; 1.  
CC PROSITE; PS00764; ENDONUCLEASE\_III\_1; 1.  
CC PROSITE; PS01155; ENDONUCLEASE\_III\_2; 1.  
CC Hydrolase; Nuclease; Endonuclease; Multifunctional enzyme; DNA repair;  
KW Glycosidase; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.  
FT METAL 198 198 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 205 205 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 208 208 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 214 214 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
SQ SEQUENCE 245 AA; 27030 MW; 2B6D16195DD090DE CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
DB 91 GFYRN 95  
RESULT 5  
ID UVEN MICLU STANDARD; PRT; 279 AA.  
AC P46303;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ultraviolet N-glycosylase/AP lyase (UV-endonuclease) (Pyrimidine dimer  
DE glycosylase).  
GN PDG.  
OS Micrococcus luteus (Micrococcus lysodeikticus).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococciaceae; Micrococaceae; Micrococcaceae; Micrococcus.  
RN [1]\_TaxID=1270;  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-35.  
RC STRAIN=ATCC 4698;  
RX MEDLINE=96007490; PubMed=7559510;  
RA Pietsen C.E., Prince M.A., Augustine M.L., Dodson M.L., Lloyd R.S.;  
RT "Purification and cloning of Micrococcus luteus ultraviolet  
RT endonuclease, an N-glycosylase/abasic lyase that proceeds via an  
RT imino enzyme-DNA intermediate.";  
RL J. Biol. Chem. 270:23475-23484(1995).  
CC -1- FUNCTION: HAS BOTH, AN APURINIC AND/OR APYRIMIDINIC ENDONUCLEASE  
CC ACTIVITY AND A DNA N-GLYCOSYLASE ACTIVITY. INITIATES REPAIR AT  
CC CIS-SYN PYRIMIDINE DIMERS. PROCEEDS VIA AN IMINO ENZYME:DNA  
CC INTERMEDIATE.  
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN  
CC CODONS FOR GLY-268 AND ALA-270. TWO FORMS OF 31 KDA AND 32 KDA  
CC HAVE BEEN DETECTED.  
CC -1- SIMILARITY: BELONGS TO THE NTH/MUTY FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U22181; AAA86508.1; ALT\_TERM.  
CC HSSP; P20625; 2ABK.  
CC InterPro; IPR003265; Endo\_3c.  
CC InterPro; IPR004035; EndoIII\_FCL.  
CC InterPro; IPR004036; EndoIII\_HhH.  
CC InterPro; IPR003651; FeS bind.  
CC InterPro; IPR000445; HhH.  
CC InterPro; IPR003583; HhH\_1.  
CC InterPro; IPR005759; Nth.  
CC Pfam; PF00730; HhH-GPD; 1.  
CC Pfam; PF00633; HhH; 1.  
CC SMART; SM00478; ENDO3c; 1.  
CC SMART; SM00525; FES; 1.  
CC SMART; SM00278; HhH1; 1.  
CC TIGRfams; TIGR01083; nth; 1.  
CC PROSITE; PS00764; ENDONUCLEASE\_III\_1; 1.  
CC PROSITE; PS01155; ENDONUCLEASE\_III\_2; 1.  
CC Hydrolase; Nuclease; Endonuclease; DNA repair; Glycosidase;  
KW Hydrolyase; Nuclease; Endonuclease; DNA repair; Glycosidase;  
FT CHAIN 1 268 UV ENDONUCLEASE 31 kDa FORM.  
FT CHAIN 1 279 UV ENDONUCLEASE 32 kDa FORM (PROBABLE).  
FT FT

```
FT METAL 203 203 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 210 210 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 213 213 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 219 219 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 279 AA; 30469 MW; C385369A1827C005 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 95 GFYRN 99

RESULT 6
SERC_XANAC STANDARD; PRT; 361 AA.
AC Q8PLV7;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XAC1648.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos M.V.F.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -|- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonooxypyruvate + L-glutamate.
CC -|- COPACTOR: Pyridoxal phosphate (By similarity).
CC -|- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE011797; AAM36516.1; -.
CC HAMAP; MF_00160; -.
CC InterPro; IPR000192; Aminotransf.
CC Pfam; PF00266; aminotransf_5; 1.
CC
```

```
DR ProDom; PD001544; Pser aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC1; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
KW Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 196 196 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 38680 MW; 75EC2B319C493982 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 283 GFYRN 287

RESULT 7
SERC_XANCP STANDARD; PRT; 361 AA.
AC Q8PA97;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XCC1589.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos M.V.F.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -|- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonooxypyruvate + L-glutamate.
CC -|- COPACTOR: Pyridoxal phosphate (By similarity).
CC -|- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012260; AAM40884.1; -.
CC HAMAP; MF_00160; -.
CC InterPro; IPR000192; Aminotransf.
CC
```



```
DR InterPro; IPR003248; Peer aminotransf.
DR Pfam; PF00266; aminotran 5; 1.
DR ProDom; PD001544; Peer aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC 1; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS 5; 1.
KW Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 196 196 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 38759 MW; F0025FFB9BF7B65 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 283 GFYRN 287

RESULT 8
SERC_YERPE
ID SERC_YERPE STANDARD; PRT; 361 AA.
AC Q8ZG64;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR YP01389 OR Y2784.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holtroyd S., Jagsels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RA "Genome sequence of Versinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Lies P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Milles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RA "Genome sequence of Versinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonoxypropylate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; AJ414148; CAC90218.1; -.
DR EMBL; AE013881; AM86336.1; -.
DR FIR; AG0169; AG0169.
DR HAMAP; MF_00160; -. 1.
DR InterPro; IPR00192; Aminotransfv.
DR InterPro; IPR003248; Pser aminotransf.
DR Pfam; PF00266; aminotran 5; 1.
DR ProDom; PD001544; Peer aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC 1; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS 5; 1.
KW Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 197 197 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40083 MW; 8827EF1419782D88 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 283 GFYRN 287

RESULT 9
SERC_XYLFA
ID SERC_XYLFA STANDARD; PRT; 362 AA.
AC Q9PB19;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XF2326.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RA "The genome sequence of the plant pathogen Xylella fastidiosa.";
RT Nature 406:151-159(2000).
RL Nature 406:151-159(2000).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonoxypropylate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL646061; CAD14605.1; -.
CC HAMAP; MF 00160; -; 1.
CC InterPro; IPR000192; Aminotransf.
CC InterPro; IPR003248; Pser_aminotransf.
CC Pfam; PF00266; aminotran_5; 1.
CC ProDom; PD001544; Pser_aminotransf; 1.
CC TIGRFAMs; TIGR01364; serC_1; 1.
CC PROSITE; PS00595; AA_TRANSFER_CLASS_5; FALSE_NEG.
CC Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
CC Aminotransferase; Pyridoxal phosphate; Complete proteome.
CC FT BINDING 214 214 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC SEQUENCE 378 AA; 41860 MW; D4CA356D9098E97B CRC64;
CC -----
CC Query Match 93.8%; Score 30; DB 1; Length 378;
CC Best Local Similarity 100.0%; Pred. No. 25;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 3 GFYRN 7
CC |||||
CC 300 GFYRN 304
CC -----
CC RESULT 11
CC PUR2 RHIME STANDARD; PRT; 423 AA.
CC ID PUR2_RHIME
CC AC Q92RL0;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide
CC DE ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase).
CC DE PURD OR R00858 OR SMC00993.
CC OS Rhizobium melliloti (Sinorhizobium melloti).
CC GN Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
CC OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
CC OX NCBI_TaxID=382;
CC [1]
CC SEQUENCE FROM N.A.
CC RP STRAIN=1021;
CC RC MEDLINE=21396507; PubMed=11481430;
CC RA Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
CC RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
CC RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
CC RA Pohl T., Portetelle D., Puenhler A., Purnelle B., Ramsperger U.,
CC RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
CC RT "Analysis of the chromosome sequence of the legume symbiont
CC RT Sinorhizobium melliloti strain 1021.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001)
CC -|- CATABOLIC ACTIVITY: ATP + 5-phospho-D-riboseylamine + glycine = ADP
CC + phosphate + N(1)-(5-phospho-D-riboseyl)glycinamide.
CC -|- PATHWAY: De novo purine biosynthesis; second step.
CC -|- SIMILARITY: Belongs to the GARS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL591785; CAC45430.1; -.
CC HAMAP; MF 00138; -; 1.
CC InterPro; IPR000115; Gars.
CC -----

```

DR Pfam; PF01071; GARS; 1.  
 DR PFam; PF02842; GARS B; 1.  
 DR PFam; PF02843; GARS C; 1.  
 DR PFam; PF02844; GARS N; 1.  
 DR TIGRFAMs; TIGR00877; purD; 1.  
 DR PROSITE; PS00184; GARS; 1.  
 DR Purine biosynthesis; Ligase; Complete proteome.  
 SQ SEQUENCE 423 AA; 44324 MW; 5E65E13B606D204B CRC64;

Query Match 93.8%; Score 30; DB 1; Length 423;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 407 GFYRN 411

RESULT 12  
 SYW MOUSE  
 ID SYW MOUSE STANDARD; PRT; 481 AA.  
 AC P32921;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tyryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tyryptophan--trna ligase) (TrpRS).  
 DE WARS OR WRS.  
 GN WARS musculus (Mouse).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=95018226; PubMed=7932716;  
 RA Fajot B., Sarger C., Bonnet J., Garret M.;  
 RT "An alternative splicing modifies the C-terminal end of tryptophanyl-  
 trna synthetase in murine embryonic stem cells.";  
 RL J. Mol. Biol. 242:599-603(1994).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=Long;  
 CC IsoID=P32921-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Short;  
 CC IsoID=P32921-2; Sequence=VSP\_006313;  
 CC TISSUE SPECIFICITY: Isoform 2 is widely expressed, isoform 1 is  
 found only in embryonic stem cells.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X69656; CAA49347.1; -;  
 CC EMBL; X69657; CAA49348.1; -;  
 CC PIR; S50053; S50053.  
 CC MGI; 104630; Wars.  
 CC InterPro; IPR002305; trna-synt\_1b.  
 CC InterPro; IPR001412; trna-synt\_1.  
 CC InterPro; IPR002306; Trp trna-synt\_1b.  
 CC InterPro; IPR000738; WHEP-TRS.  
 CC Pfam; PF00579; trna-synt\_1b; 1.  
 CC Pfam; PF00458; WHEP-TRS; 1.  
 CC PRINTS; PR01039; TRNASYNTHTRP.  
 CC TIGRFAMs; TIGR00233; trps; 1.

DR PROSITE; PS00178; AA TRNA LIGASE\_I; 1.  
 DR PROSITE; PS00762; WHEP TRS; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Alternative splicing.  
 FT DOMAIN 23 68  
 FT SITE 168 177  
 FT SITE 353 357  
 FT VARSPLIC 476 481  
 FT /FTID=VSP\_006313.  
 SQ SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 481;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 250 GFYRN 254

RESULT 13  
 TYRO MOUSE  
 ID TYRO MOUSE STANDARD; PRT; 533 AA.  
 AC P11344;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tyrosinase precursor (EC 1.14.18.1) (Monophenol monooxygenase)  
 DE (Albino locus protein).  
 GN TYR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2J;  
 RX MEDLINE=88268910; PubMed=3134020;  
 RA Kwon B.S., Wakulchik M., Haq A.K., Halaban R., Kestler D.;  
 RT "Sequence analysis of mouse tyrosinase cDNA and the effect of  
 melanotropin on its gene expression.";  
 RL Biochem. Biophys. Res. Commun. 153:1301-1309(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Himalayan;  
 RX MEDLINE=89273644; PubMed=2567165;  
 RA Kwon B.S., Halaban R., Chintamani C.;  
 RT "Molecular basis of mouse Himalayan mutation.";  
 RL Biochem. Biophys. Res. Commun. 161:252-260(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89030636; PubMed=3141148;  
 RA Mueller G., Ruppert S., Schmid E., Schuetz G.;  
 RT "Functional analysis of alternatively spliced tyrosinase gene  
 transcripts.";  
 RL EMBO J. 7:2723-2730(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89193679; PubMed=2494997;  
 RA Terao M., Tabe L., Garattini E., Sartori D., Studer M., Mintz B.;  
 RT "Isolation and characterization of variant cDNAs encoding mouse  
 tyrosinase.";  
 RL Biochem. Biophys. Res. Commun. 159:848-853(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J;  
 RA Yamamoto H., Takeuchi S., Kudo T., Makino K., Nakata A., Shinoda T.,  
 RA Takeuchi T.;  
 RT "Cloning and sequencing of mouse tyrosinase cDNA.";  
 RL Jpn. J. Genet. 62:271-274(1987).  
 RN [6]  
 RP SEQUENCE OF 1-273 FROM N.A.  
 RX MEDLINE=90212084; PubMed=2517217;



```
Query Match          93.8%; Score 30; DB 1; Length 824;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
      |||||
DB      206 GFYRN 210

RESULT 15
XPC_MOUSE STANDARD; PRT; 900 AA.
AC P51612; P54732;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA-repair protein complementing XP-C cells homolog (Xeroderma
DE pigmentosum group C complementing protein homolog) (P125).
GN XPC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96184849; PubMed=8604333;
RA Li L., Peterson C., Legerski R.;
RT "Sequence of the mouse XPC cDNA and genomic structure of the human
RT XPC gene.";
RL Nucleic Acids Res. 24:1026-1028 (1996).
RN [2]
RP SEQUENCE OF 28-587 FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=95405469; PubMed=7675084;
RA Sands A.T., Abuin A., Sanchez A., Conti C.J., Bradley A.;
RT "High susceptibility to ultraviolet-induced carcinogenesis in mice
RT lacking XPC.";
RL Nature 377:162-165 (1995).
CC -!- FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA
CC DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO
CC ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.
CC -!- SUBUNIT: HETERODIMER OF A 125 Kda SUBUNIT (P125) AND OF A
CC 58 Kda SUBUNIT (P58).
CC -!- SUBCELLULAR LOCATION: Nuclear (probable).
CC -!- SIMILARITY: SOME, TO YEAST RAD4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U27398; AAC52500.1; --
DR EMBL; U40005; AAA82720.1; --
DR PIR; S70630; S70630.
DR MGD; MGI:103557; Xpc.
DR GO; GO:0006289; P:nucleotide-excision repair; IMP.
DR InterPro; IPR004583; Rad4.
DR Pfam; PF03835; Rad4; 1.
DR TIGRfams; TIGR00605; rad4; 1.
KW DNA repair; DNA-binding; Nuclear protein.
FT CONFLICT 28 30 AVN -> CSD (IN REF. 2).
FT CONFLICT 53 53 S -> L (IN REF. 2).
FT CONFLICT 67 67 L -> F (IN REF. 2).
FT CONFLICT 70 70 L -> S (IN REF. 2).
FT CONFLICT 134 135 RG -> TP (IN REF. 2).
FT CONFLICT 165 170 EQVENM -> GVHEDT (IN REF. 2).
FT CONFLICT 181 181 S -> N (IN REF. 2).
FT CONFLICT 187 187 S -> N (IN REF. 2).
FT CONFLICT 190 190 R -> S (IN REF. 2).

Query Match          93.8%; Score 30; DB 1; Length 900;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
      |||||
DB      182 GFYRN 186

Search completed: December 8, 2003, 09:14:13
Job time : 2.83838 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 7.35354 Seconds  
(without alignments)  
245.646 Million cell updates/sec

Title: US-09-498-556C-79  
Perfect score: 32  
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	77	Q8X3Q1	Q8x3q1 escherichia
2	30	93.8	88	Q8FIE2	Q8fie2 escherichia
3	30	93.8	91	Q16938	Q16938 ancylostoma
4	30	93.8	102	Q962V8	Q962v8 ancylostoma
5	30	93.8	110	Q9KYV8	Q9kyv8 streptomyce
6	30	93.8	134	P74345	P74345 synchocyst
7	30	93.8	154	Q8ITP8	Q8itp8 oesophagost
8	30	93.8	158	Q44490	Q44490 anabaena va
9	30	93.8	172	Q9A2A6	Q9a2a6 caulobacter
10	30	93.8	178	Q9TAJ9	Q9taj9 cafeteria r
11	30	93.8	184	Q9VGI2	Q9vgi2 drosophila
12	30	93.8	184	Q96723	Q96723 drosophila
13	30	93.8	191	Q8L803	Q8l803 triticum ae
14	30	93.8	216	Q8TNS2	Q8tns2 methanosarc
15	30	93.8	222	Q8DIE9	Q8die9 synchococc
16	30	93.8	223	Q8YQ65	Q8yq65 anabaena sp

17	30	93.8	228	Q8GMR5	Q8gmr5 synchococc
18	30	93.8	230	Q8U2B5	Q8u2b5 grapevine f
19	30	93.8	237	Q8HPV4	Q8hvp4 halobacteri
20	30	93.8	260	Q8NTL4	Q8ntl4 corynebacte
21	30	93.8	264	Q8FSU2	Q8fsu2 corynebacte
22	30	93.8	266	Q8ZTI9	Q8zt19 pyrobaculum
23	30	93.8	302	Q8C2V6	Q8c2v6 mus musculu
24	30	93.8	325	Q9BKX3	Q9bkx3 lucilia cup
25	30	93.8	354	Q9VEY0	Q9vey0 drosophila
26	30	93.8	361	Q8PLX7	Q8ply7 xanthomonas
27	30	93.8	361	Q8PA97	Q8pa97 xanthomonas
28	30	93.8	361	Q8ZGB4	Q8zgb4 yersinia pe
29	30	93.8	362	Q9PBI9	Q9pb19 xylella fas
30	30	93.8	378	Q8Y0Z0	Q8y0z0 ralstonia s
31	30	93.8	408	Q01482	Q01482 caenorhabdi
32	30	93.8	415	Q9U153	Q9u153 leishmania
33	30	93.8	465	Q9C8U9	Q9c8j9 arabidopsis
34	30	93.8	470	Q8EDG2	Q8edg2 shewanella
35	30	93.8	472	Q02868	Q02868 saccharomyc
36	30	93.8	475	Q9P6K0	Q9p6k0 schizosacch
37	30	93.8	475	Q9DC65	Q9dc65 mus musculu
38	30	93.8	481	Q99J58	Q99j58 mus musculu
39	30	93.8	511	Q8GRX1	Q8grx1 arabidopsis
40	30	93.8	533	Q91XK0	Q91xk0 mus musculu
41	30	93.8	554	Q8ZZL2	Q8zzl2 pyrobaculum
42	30	93.8	596	Q86468	Q86468 rhodobacter
43	30	93.8	600	Q9NL27	Q9nl27 ciona intes
44	30	93.8	611	Q9L775	Q9l775 arabidopsis
45	30	93.8	611	Q8GZT8	Q8gzt8 arabidopsis

## ALIGNMENTS

## RESULT 1

Q8X3Q1 PRELIMINARY; PRT; 77 AA.  
AC Q8X3Q1;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE Hypothetical protein z1836.  
GN Z1836.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocous K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
DR EMBL; AE005327; AAG55934.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 77 AA; 8087 MW; F85D56712A58545E CRC64;

Query Match 93.8%; Score 30; DB 16; Length 77;  
Best Local Similarity 100.0%; Pred.No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 15 GFYRN 19

## RESULT 2

```
Q8FIE2
ID Q8FIE2 PRELIMINARY; PRT; 88 AA.
AC Q8FIE2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN C1484
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016759; AAN79953.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 9179 MW; 6709ABAC8259B8B CRC64;

Query Match 93.8%; Score 30; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 15 GFYRN 19

RESULT 3
Q16938 PRELIMINARY; PRT; 91 AA.
ID Q16938
AC Q16938
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anti-coagulant protein C2 precursor (Fragment).
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98298519; PubMed=9634780;
RA Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,
RA van den Brande I., Gansemans Y.G., Lauwereys M.J., Vlasuk G.P.,
RA Stanssens P.E.;
RT "Surface expression and ligand-based selection of cDNAs fused to
RT filamentous phage gene VI.";
RL Biotechnology 13:378-382(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312555; PubMed=8700900;
RA Stanssens P., Bergum P.W., Gansemans Y., Jaspers L., Laroche Y.,
RA Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,
RA Lasters I., Vlasuk G.P.;
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).
DR EMBL; U30793; AAC47080.1; -.
DR HSSP; P56682; 1CCV.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal.
FT SIGNAL 1 1
```

```
FT SIGNAL <1 7 POTENTIAL.
FT CHAIN 8 91 ANTI-COAGULANT PROTEIN C2.
SQ SEQUENCE 91 AA; 10358 MW; EC811CB4597C24DA CRC64;

Query Match 93.8%; Score 30; DB 5; Length 91;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 62 GFYRN 66

RESULT 4
Q962V8 PRELIMINARY; PRT; 102 AA.
ID Q962V8
AC Q962V8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anticoagulant peptide-1 precursor.
GN ACEAP-1.
OS Ancylostoma ceylanicum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=53326;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison L.M., Cappello M.;
RT "Cloning of the major factor Xa inhibitor (AceAP-1) from Ancylostoma
RT ceylanicum";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399710; AAK81733.1; -.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 102 ANTI-COAGULANT PEPTIDE-1.
SQ SEQUENCE 102 AA; 11834 MW; 315722980EF723E7 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 71 GFYRN 75

RESULT 5
Q9KYV8 PRELIMINARY; PRT; 110 AA.
ID Q9KYV8
AC Q9KYV8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SCO3187.
GN SCO3187 OR SCE22.04.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
```



RA Rabbinitzsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2)." (2002).  
 RT Nature 417:141-147 (2002).  
 DR EMBL; AL939115; CAB90971.1; -;  
 SQ SEQUENCE 110 AA; 10831 MW; 0893F0F52B8EA8FA CRC64;  
 Query Match 93.8%; Score 30; DB 16; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 51 GFYRN 55  
 RESULT 6  
 P74345 PRELIMINARY; PRT; 134 AA.  
 ID P74345  
 AC P74345  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein slr1628.  
 GN SLR1628.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsumoto A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136 (1996).  
 DR EMBL; D90914; BAA18439.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 134 AA; 14545 MW; 7E5414E80FB43D10 CRC64;  
 Query Match 93.8%; Score 30; DB 16; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 27 GFYRN 31  
 RESULT 7  
 Q81TP8 PRELIMINARY; PRT; 154 AA.  
 ID Q81TP8  
 AC Q81TP8  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative trypsin-like inhibitor protein precursor.  
 GN MCRP.  
 OS Oesophagostomum dentatum.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Strongyloidea; Chabertiidae; Oesophagostomum.  
 OX NCBI\_TaxID=61180;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boag P.R., Ranganathan S., Newton S.E., Gasser R.B.;

RT "Identification of a male-specific nematode protein with two trypsin  
 RT like inhibitor domains.";  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF399936; AAN32637.1; -;  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 154 PUTATIVE TRYPSIN-LIKE INHIBITOR PROTEIN.  
 SQ SEQUENCE 154 AA; 16564 MW; A7F566E3957DA819 CRC64;  
 Query Match 93.8%; Score 30; DB 5; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 87;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 64 GFYRN 68  
 RESULT 8  
 Q44490 PRELIMINARY; PRT; 158 AA.  
 ID Q44490  
 AC Q44490  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Orf2.  
 OS Anabaena variabilis.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=1172;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 29413;  
 RX MEDLINE=96016168; PubMed=7568132;  
 RA Thiel T., Lyons E.M., Erker J.C., Ernst A.;  
 RT "A second nitrogenase in vegetative cells of a heterocyst-forming  
 RT cyanobacterium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29413;  
 RA Thiel T., Lyons E.M., Erker J.C.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U49859; AAA93026.1; -;  
 DR InterPro; IPR004952; DUF269.  
 DR Pfam; PF03270; DUF269; 1  
 SQ SEQUENCE 158 AA; 17791 MW; D157EBF59C36FEED CRC64;  
 Query Match 93.8%; Score 30; DB 2; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 35 GFYRN 39  
 RESULT 9  
 Q9A2A6 PRELIMINARY; PRT; 172 AA.  
 ID Q9A2A6  
 AC Q9A2A6  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein CC3660.  
 GN CC3660.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 MEDLINE=21173698; PubMed=11259647;

```

RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Winn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Emolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RA "Complete genome sequence of Caulobacter crescentus."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AEO06024; AAK25622.1; -.
DR TIGR; CC3660; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 172 AA; 18425 MW; 5DD52E712F406F6A CRC64;

Query Match 93.8%; Score 30; DB 16; Length 172;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 73 GFYRN 77

RESULT 10
Q9TAJ9 PRELIMINARY; PRT; 178 AA.
AC Q9TAJ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Orf178.
GN Orf178.
OS Cafeteria roenbergensis.
OC Eukaryota; stramenopiles; Bicosoecida; Cafeteria.
OX NCBI_TaxID=33653;
RN [1]
RP SEQUENCE FROM N.A.
RA Burger G.;
RT "The mitochondrial genome of Cafeteria roenbergensis."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF193903; AAF05787.1; -.
KW Mitochondrion.
SQ SEQUENCE 178 AA; 21418 MW; 00F69B2FC8461362 CRC64;

Query Match 93.8%; Score 30; DB 8; Length 178;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 109 GFYRN 113

RESULT 11
Q9VGJ2 PRELIMINARY; PRT; 184 AA.
AC Q9VGJ2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE I-T protein.
GN I-T OR CGI4719.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Durbin K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003692; AAF54687.1; -.
DR FlyBase; FBgn0025821; I-T.
SQ SEQUENCE 184 AA; 21018 MW; 97861392D5B77069 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 31 GFYRN 35

RESULT 12
Q96723 PRELIMINARY; PRT; 184 AA.
AC Q96723;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Inhibitor-T protein.
GN I-T OR INHIBITOR-T OR CGI4719.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=CANTON S;
RX MEDLINE=99037736; PubMed=9821974;
RA Helms N.R., Vergidou C., Gaskell T., Cohen P.T.W.;
RT "Characterisation of a novel Drosophila melanogaster testis specific
RT PPI inhibitor related to mammalian inhibitor-2: identification of the
RT site of interaction with PPI."
RL FEBS Lett. 438:131-136(1998).

```

```
DR EMBL; AJ006867; CAA07278.1; -.
DR FlyBase; FBgn0025821; I-t.
FT VARIANT 21 21 A -> T.
FT VARIANT 183 183 D -> G.
SQ SEQUENCE 184 AA; 20988 MW; D697139E7271A7B2 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 31 GFYRN 35

RESULT 13
Q8L803 PRELIMINARY; PRT; 191 AA.
AC Q8L803;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative plastid ribosomal protein CL9.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Z.-X., Laroche A., Gaudet D.;
RT "Triticum aestivum putative plastid ribosomal protein CL9 gene.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV123421; RA92711.1; -.
DR InterPro; IPR000244; Ribosomal_L9.
DR Pfam; PF03948; Ribosomal_L9_C; 1.
DR Pfam; PF01281; Ribosomal_L9_N; 1.
DR TIGRFAMs; TIGR00158; L9; 1.
DR PROSITE; PS00651; RIBOSOMAL_L9; 1.
KW Ribosomal protein.
SQ SEQUENCE 191 AA; 21582 MW; DB6D7E72B812A34 CRC64;

Query Match 93.8%; Score 30; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 67 GFYRN 71

RESULT 14
Q8TNS2 PRELIMINARY; PRT; 216 AA.
AC Q8TNS2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Trk potassium uptake system protein.
GN TRKA OR MA2210.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels-Thomann N., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,

RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010907; RA05606.1; -.
DR InterPro; IPR003148; TrkA_N; 1.
DR Pfam; PF02254; TrkA_N; 1.
KW Complete proteome.
SQ SEQUENCE 216 AA; 24016 MW; 9E64363A9CE9FBF9 CRC64;

Query Match 93.8%; Score 30; DB 17; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 175 GFYRN 179

RESULT 15
Q8DIE9 PRELIMINARY; PRT; 222 AA.
AC Q8DIE9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Endonuclease III.
GN TLL1641.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005374; BAC09193.1; -.
KW Complete proteome.
SQ SEQUENCE 222 AA; 25106 MW; 1BE6A1B5AA4ED777 CRC64;

Query Match 93.8%; Score 30; DB 16; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 88 GFYRN 92

Search completed: December 8, 2003, 09:18:39
Job time : 9.35353 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 2.89899 Seconds  
(without alignments)  
232.212 Million cell updates/sec

Title: US-09-498-556C-79  
Perfect score: 32  
Sequence: 1 XXGFYRN 77

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	77	B85684	unknown protein en
2	30	93.8	134	S76180	hypothetical prote
3	30	93.8	172	B87703	conserved hypothet
4	30	93.8	175	S73496	hypoxanthine-guani
5	30	93.8	175	F64250	hypoxanthine-guani
6	30	93.8	223	AC2302	endonuclease III l
7	30	93.8	226	C70790	probable Endonucle
8	30	93.8	237	G84299	hypothetical prote
9	30	93.8	361	AG0169	phosphoserine tran
10	30	93.8	362	C82572	phosphoserine tran
11	30	93.8	408	T25524	hypothetical prote
12	30	93.8	415	T46716	hypothetical prote
13	30	93.8	465	A96553	probable myrosinas
14	30	93.8	481	S50053	cryptophan-tRNA li
15	30	93.8	533	1 YRMSCS	monophenol monooxy
16	30	93.8	747	T16274	hypothetical prote
17	30	93.8	824	S40937	hypothetical prote
18	30	93.8	900	S70630	xeroderma pigmento
19	30	93.8	940	S44345	AcrB/AcrD/AcrF fam
20	30	93.8	1041	C87645	probable transport
21	27	84.4	1564	S55517	hypothetical prote
22	27	84.4	36	H70251	ribosomal protein
23	27	84.4	60	F70187	hypothetical prote
24	27	84.4	199	T39498	hypothetical prote
25	27	84.4	201	A86636	purine-nucleoside
26	27	84.4	234	A84593	hypothetical prote
27	27	84.4	258	T22092	conserved hypothet
28	27	84.4	275	A69413	hypothetical prote
29	27	84.4	299	F64491	hypothetical prote

30	27	84.4	318	2	B36972	aad 5'-region hypo
31	27	84.4	318	2	S33433	hypothetical prote
32	27	84.4	340	2	D69414	hypothetical prote
33	27	84.4	340	2	C75004	hypothetical prote
34	27	84.4	353	2	E64376	endonuclease III -
35	27	84.4	425	2	F69193	acetyl-CoA synthet
36	27	84.4	426	2	F69994	uncharacterized co
37	27	84.4	461	2	E95152	v-type sodium ATP
38	27	84.4	463	2	H85745	unknown protein en
39	27	84.4	463	2	C90873	hypothetical prote
40	27	84.4	465	2	D64881	ycjX protein - Esc
41	27	84.4	465	2	D90866	probable enzyme (i
42	27	84.4	465	2	E85752	probable EC 2.1 en
43	27	84.4	465	2	AC0659	probable ATP-bindi
44	27	84.4	471	1	A41706	tryptophan-tRNA li
45	27	84.4	471	1	S62590	peptidyl-prolyl ci

#### ALIGNMENTS

##### RESULT 1

B85684

unknown protein encoded by prophage CP-933C [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: B85684

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85684

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-77 <STO>

A:Cross-references: GB:AE005174; NID:G12514756; PIDN:AA055934.1; GSPDB:GN00145; UWGP:Z18

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z1836

Query Match 93.8%; Score 30; DB 2; Length 77;

Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

DB 15 GFYRN 19

##### RESULT 2

S76180

hypothetical protein - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S76180

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.,

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasudr

DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76180

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-134 <KAN>

A:Cross-references: EMBL:D90914; GB:AB001339; NID:G1653477; PIDN:BAAL8439.1; PID:dl0191

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 100.0%; Pred. No. 17;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 27 GFYRN 31

## RESULT 3

B87703 conserved hypothetical protein CC3660 [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C;Accession: B87703  
R;Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: B87703  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-172 <STO>  
A;Cross-references: GB:AE005673; NID:gl3425418; PIDN:AAK25622.1; GSPDB:GN00148  
C;Genetics:  
A;Gene: CC3660

Query Match 93.8%; Score 30; DB 2; Length 172;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 73 GFYRN 77

## RESULT 4

S73496 hypoxanthine-guanine phosphoribosyltransferase hpt - Mycoplasma pneumoniae (strain ATCC  
N;Alternate names: hypothetical protein K05\_orf175  
C;Species: Mycoplasma pneumoniae  
A;Variety: ATCC 29342  
C;Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C;Accession: S73496  
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A;Reference number: S73327; MUID:97105885; PMID:8948633  
A;Accession: S73496  
A;Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-175 <HM>  
A;Cross-references: EMBL:AE000018; GB:U00089; NID:gl673827; PIDN:AAB95818.1; PID:gl67383  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C;Genetics:  
A;Gene: hpt  
A;Genetic code: SGC3  
C;Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 93.8%; Score 30; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 159 GFYRN 163

## RESULT 5

F64250 hypoxanthine-guanine phosphoribosyltransferase - Mycoplasma genitalium  
C;Species: Mycoplasma genitalium  
C;Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-Dec-1999  
C;Accession: F64250

R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.;  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A;Title: The minimal gene complement of Mycoplasma genitalium.  
A;Reference number: A64200; MUID:96026346; PMID:7569993  
A;Accession: F64250  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-175 <TIGR>  
A;Cross-references: GB:U39733; GB:L43967; NID:gl046177; PIDN:AAB01648.1; PID:gl046178; T:  
A;Experimental source: strain G-37  
C;Genetics:  
A;Genetic code: SGC3  
C;Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 93.8%; Score 30; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 159 GFYRN 163

## RESULT 6

AC2302 endonuclease III [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AC2302  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AC2302  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-223 <GUR>  
A;Cross-references: GB:BA000019; PIDN:BAW75669.1; PID:gl7133104; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: nth  
C;Superfamily: apurinic/aprymidinic endonuclease III

Query Match 93.8%; Score 30; DB 2; Length 223;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 89 GFYRN 93

## RESULT 7

C70790 probable Endonuclease III - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C;Accession: C70790  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: C70790  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-226 <COL>

A;Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA17996.1; PID:g296009  
A;Experimental source: strain H37Rv  
C;Genetics:

A;Gene: nth  
C;Superfamily: apurinic/apyrimidinic endonuclease III  
C;Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein  
F;179,186,189,195/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 93.8%; Score 30; DB 2; Length 226;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 72 GFYRN 76

## RESULT 8

G84299  
hypothetical protein Vng1457c [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-feb-2001 #text\_change 02-Feb-2001  
C;Accession: G84299  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebbhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: G84299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-237 <STO>  
A;Cross-references: GB:AE004437; NID:g10580953; PIDN:AAG19763.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: VNG1457C

Query Match 93.8%; Score 30; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 191 GFYRN 195

## RESULT 9

AG0169  
phosphoserine transaminase (EC 2.6.1.52) [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AG0169  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AG0169  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-361 <KUR>  
A;Cross-references: GB:ALS90842; PIDN:CAC90218.1; PID:g15979438; GSPDB:GN00175  
C;Genetics:  
A;Gene: serC  
C;Superfamily: phosphoserine aminotransferase  
C;Keywords: aminotransferase

Query Match 93.8%; Score 30; DB 2; Length 361;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 283 GFYRN 287

## RESULT 10

C82572  
phosphoserine aminotransferase XF2326 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C;Accession: C82572  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number AS9328 below  
A;Accession: C82572  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-362 <SIM>  
A;Cross-references: GB:AE004043; GB:AE003849; NID:g9107486; PIDN:AAF85125.1; GSPDB:GN001  
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, I  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, I  
A;Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaas  
A;Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsuchioka, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; V  
A;Reference number: AS9328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF2326  
C;Superfamily: phosphoserine aminotransferase

Query Match 93.8%; Score 30; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 284 GFYRN 288

## RESULT 11

T25524  
hypothetical protein C06A5.9 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T25524  
R;Davidson, S.; Wohldmann, P.  
submitted to the EMBL Data Library, April 1997  
A;Description: The sequence of C. elegans cosmid C06A5.  
A;Reference number: Z20044  
A;Accession: T25524  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-408 <DAV>  
A;Cross-references: EMBL:U97193; PIDN:AAB52442.1; GSPDB:GN00019; CESP:C06A5.9  
A;Experimental source: strain Bristol N2; clone C06A5  
C;Genetics:  
A;Gene: CESP:C06A5.9  
A;Map position: 1  
A;Introns: 3/1; 27/3; 86/1; 154/1; 201/3; 240/3; 268/3; 297/3  
Query Match 93.8%; Score 30; DB 2; Length 408;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

C;Date: 07-May-1995 #sequence\_revision 10-Nov-1995 #text\_change 03-Jun-2002  
C;Accession: S50053; S50052; I49391; S31461; S31462  
R;Fajot, B.; Sarger, C.; Bonnet, J.; Garret, M.  
J. Mol. Biol. 242, 599-603, 1994  
A;Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthetase  
A;Reference number: S50052; MUID:95018226; PMID:7932716  
A;Accession: S50053  
A;Molecule type: mRNA  
A;Residues: 1-481 <PAJ>  
A;Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438  
A;Genetics: LSF  
A;Note: Intron position was determined by sequencing of genomic DNA  
A;Accession: S50052  
A;Molecule type: mRNA  
A;Residues: 1-475 <PAW>  
A;Cross-references: EMBL:X69656; NID:G55435; PIDN:CAA49347.1; PID:G55436  
A;Genetics: SSF  
R;Kisselev, L.L.  
Biochimie 75, 1027-1039, 1993  
A;Title: Mammalian tryptophanyl-tRNA synthetases.  
A;Reference number: I49391; MUID:94257729; PMID:7515282  
A;Accession: I49391  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-481 <RES>  
A;Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438  
C;Genetics: <LSF>  
A;Gene: WRS  
A;Introns: 475/2  
A;Note: the list of introns may be incomplete; clone W13  
C;Genetics: <SSP>  
A;Gene: WRS  
A;Note: clone S5  
C;Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C;Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis  
F;1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted  
F;1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted  
F;23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 93.8%; Score 30; DB 2; Length 481;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 250 GFYRN 254

RESULT 15  
YRMSCS  
monophenol monooxygenase (EC 1.14.18.1) precursor [validated] - mouse  
N;Alternate names: cresolase; monophenol oxidase; phenolase; tyrosinase  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 20-Apr-2000  
C;Accession: A27711; A60778; A32429; S01170; S02278; S15753; I49736  
R;Kwon, B.S.; Wakulchik, M.; Haq, A.K.; Halaban, R.; Kestler, D.  
Biochem. Biophys. Res. Commun. 153, 1301-1309, 1988  
A;Title: Sequence analysis of mouse tyrosinase cDNA and the effect of melanotropin on its  
A;Reference number: A27711; MUID:88268910; PMID:3134020  
A;Accession: A27711  
A;Molecule type: mRNA  
A;Residues: 1-533 <KWO>  
A;Cross-references: GB:M20234; NID:G202247; PIDN:AAA40516.1; PID:G202248  
A;Experimental source: Cloudman S-91 melanoma cells  
R;Kwon, B.S.; Haq, A.K.; Wakulchik, M.; Kestler, D.; Barton, D.E.; Francke, U.; Lamoreux, J. Invest. Dermatol. 93, 589-594, 1989  
A;Title: Isolation, chromosomal mapping, and expression of the mouse tyrosinase gene.  
A;Reference number: A60778; MUID:90010220; PMID:2507645  
A;Accession: A60778  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-273 <KW2>  
A;Experimental source: BALB/c

Query Match 93.8%; Score 30; DB 2; Length 415;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 325 GFYRN 329

RESULT 13  
A96553  
probable myrosinase precursor 53323-50499 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C;Accession: A96553  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Liu, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: A96553  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-465 <STO>  
A;Cross-references: GB:A5005173; NID:G10092358; PIDN:RAG12767.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: FSD21.17  
A;Map position: 1  
C;Superfamily: Agrobacterium beta-glucosidase

Query Match 93.8%; Score 30; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 331 GFYRN 335

RESULT 14  
S50053  
tryptophan-tRNA ligase (EC 6.1.1.2) alpha-2 chain - mouse  
C;Species: Mus musculus (house mouse)

C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C;Accession: T46716  
R;Volckaert, G.; Ivens, A.C.; Lawson, D.; Ouail, M.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, December 1999  
A;Reference number: Z23137  
A;Accession: T46716  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-415 <VOL>  
A;Cross-references: EMBL:AL121861; PIDN:CAB58381.1  
A;Experimental source: strain Friedlin  
C;Genetics:  
A;Note: L4326.05

Query Match 93.8%; Score 30; DB 2; Length 415;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 325 GFYRN 329

hypothetical protein L4326.05 [imported] - Leishmania major  
C;Species: Leishmania major  
C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C;Accession: T46716  
R;Volckaert, G.; Ivens, A.C.; Lawson, D.; Ouail, M.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, December 1999  
A;Reference number: Z23137  
A;Accession: T46716  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-415 <VOL>  
A;Cross-references: EMBL:AL121861; PIDN:CAB58381.1  
A;Experimental source: strain Friedlin  
C;Genetics:  
A;Note: L4326.05

Query Match 93.8%; Score 30; DB 2; Length 415;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 325 GFYRN 329



R;Terao, M.; Tabe, L.; Garattini, E.; Sartori, D.; Studer, M.; Mintz, B.  
 Biochem. Biophys. Res. Commun. 159, 848-853, 1989  
 A:Title: Isolation and characterization of variant cDNAs encoding mouse tyrosinase.  
 A:Reference number: A32429; MUID:89193679; PMID:2494997  
 A:Accession: A32429  
 A:Molecule type: mRNA  
 A:Residues: 1-102, 'C', 104-345, 'G', 347-533 <TER>  
 A:Cross-references: GB:M24560; NID:G202249; PIDN:AAA40517.1; PID:G202250  
 A:Accession: B32429  
 A:Molecule type: mRNA  
 A:Residues: 1-77, 155-345, 'G', 347-533 <TE2>  
 A:Cross-references: GB:M24560  
 A:Experimental source: B16 melanoma cells  
 R;Mueller, G.; Ruppert, S.; Schmid, E.; Schuetz, G.  
 EMBO J. 7, 2723-2730, 1988  
 A:Title: Functional analysis of alternatively spliced tyrosinase gene transcripts.  
 A:Reference number: S01170; MUID:89030636; PMID:3141148  
 A:Accession: S01170  
 A:Molecule type: mRNA  
 A:Residues: 1-102, 'C', 104-263, 'I', 265-345, 'G', 347-533 <MUE>  
 A:Cross-references: GB:X12782; NID:G55061; PIDN:CAA31273.1; PID:G55062  
 R;Yamamoto, H.; Takeuchi, S.; Kudo, T.; Makino, K.; Nakata, A.; Shinoda, T.; Takeuchi, T.  
 Jpn. J. Genet. 62, 271-274, 1987  
 A:Title: Cloning and sequencing of mouse tyrosinase cDNA.  
 A:Reference number: S02278  
 A:Accession: S02278  
 A:Molecule type: mRNA  
 A:Residues: 1-102, 'C', 104-263, 'I', 265-345, 'G', 347-448 <YAM>  
 A:Cross-references: EMBL:X12782  
 A:Note: part of this sequence was confirmed by protein sequencing  
 R;Shibahara, S.; Okinaga, S.; Tomita, Y.; Takeda, A.; Yamamoto, H.; Sato, M.; Takeuchi, R.  
 Eur. J. Biochem. 189, 455-461, 1990  
 A:Title: A point mutation in the tyrosinase gene of BALB/c albino mouse causing the cyst  
 A:Reference number: S15753; MUID:90249393; PMID:2110899  
 A:Accession: S15753  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-13 <SHI>  
 A:Cross-references: EMBL:X51743; NID:G55057; PIDN:CAA36033.1; PID:G55058  
 A:Experimental source: strain BALB/c  
 R;Kwon, B.S.; Halaban, R.; Chintamaneni, C.  
 Biochem. Biophys. Res. Commun. 161, 252-260, 1989  
 A:Title: Molecular basis of mouse Himalayan mutation.  
 A:Reference number: I49736; MUID:89273644; PMID:2567165  
 A:Accession: I49736  
 A:Status: preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-39, 'I', 41-102, 'C', 104-196, 'Q', 198-345, 'G', 347-419, 'R', 421-533 <RES>  
 A:Cross-references: GB:M26729; NID:G193845; PIDN:AAA37806.1; PID:G309296  
 C:Comment: This cell-specific oxidase is a glycoprotein containing two Cu per enzyme; it  
 reactions in the formation of pigments such as melanins and other polyphenolic compound  
 C:Genetics:  
 A:Gene: Tyrl  
 A:Map position: 7  
 C:Superfamily: monophenol monooxygenase  
 C:Keywords: albinism; alternative splicing; copper; glycoprotein; melanin biosynthesis;  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-533/Product: monophenol monooxygenase #status predicted <MAT>  
 F:474-497/Domain: transmembrane #status predicted <TM>  
 F:86,111,161,230,337,371/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.8%; Score 30; DB 1; Length 533;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GYRN 7  
 |||||  
 Db 458 GYRN 462

Search completed: December 8, 2003, 09:19:32  
 Job time : 4.89899 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 1.85859 Seconds  
(without alignments)  
91.060 Million cell updates/sec

Title: US-09-498-556C-85

Perfect score: 4  
Sequence: 1 XXXX 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/aaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/aaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/aaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/aaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/aaa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	1	US-07-820-154A-12
2	0	0.0	1	1	US-07-791-213D-24
3	0	0.0	1	1	US-07-791-213D-40
4	0	0.0	1	1	US-08-174-365A-57
5	0	0.0	1	1	US-07-789-913-23
6	0	0.0	1	1	US-07-789-913-25
7	0	0.0	1	1	US-08-049-794-23
8	0	0.0	1	1	US-08-049-794-25
9	0	0.0	1	1	US-08-433-037-12
10	0	0.0	1	1	US-08-448-606-4
11	0	0.0	1	1	US-07-869-933-16
12	0	0.0	1	1	US-08-293-150A-24
13	0	0.0	1	1	US-08-293-150A-40
14	0	0.0	1	1	US-08-496-847-23
15	0	0.0	1	1	US-08-496-847-25
16	0	0.0	1	2	US-08-742-774-23
17	0	0.0	1	2	US-08-742-774-25
18	0	0.0	1	2	US-08-675-354-23
19	0	0.0	1	2	US-08-675-354-25
20	0	0.0	1	2	US-08-097-554A-12
21	0	0.0	1	2	US-08-965-918-23
22	0	0.0	1	2	US-08-965-918-25
23	0	0.0	1	2	US-09-138-439-23
24	0	0.0	1	2	US-09-138-439-25
25	0	0.0	1	3	US-08-480-640A-12
26	0	0.0	1	3	US-08-613-400A-23
27	0	0.0	1	3	US-08-613-400A-25

28 0 0.0 1 3 US-08-801-092-10 Sequence 10, Appl  
29 0 0.0 1 3 US-08-801-092-17 Sequence 17, Appl  
30 0 0.0 1 3 US-08-801-092-24 Sequence 24, Appl  
31 0 0.0 1 3 US-08-801-092-31 Sequence 31, Appl  
32 0 0.0 1 3 US-08-801-092-38 Sequence 38, Appl  
33 0 0.0 1 3 US-08-801-092-45 Sequence 45, Appl  
34 0 0.0 1 3 US-09-298-017-23 Sequence 23, Appl  
35 0 0.0 1 3 US-09-298-017-25 Sequence 25, Appl  
36 0 0.0 1 3 US-08-295-802-12 Sequence 12, Appl  
37 0 0.0 1 3 US-09-392-979A-23 Sequence 23, Appl  
38 0 0.0 1 3 US-09-392-979A-25 Sequence 25, Appl  
39 0 0.0 1 3 US-09-103-663-16 Sequence 16, Appl  
40 0 0.0 1 3 US-08-488-237A-12 Sequence 12, Appl  
41 0 0.0 1 3 US-09-117-927-5 Sequence 5, Appl  
42 0 0.0 1 4 US-08-375-992A-12 Sequence 12, Appl  
43 0 0.0 1 4 US-09-315-113-10 Sequence 10, Appl  
44 0 0.0 1 4 US-09-315-113-17 Sequence 17, Appl  
45 0 0.0 1 4 US-09-315-113-24 Sequence 24, Appl

#### ALIGNMENTS

RESULT 1  
US-07-820-154A-12  
; Sequence 12, Application US/07820154A  
; Patent No. 5382425  
; GENERAL INFORMATION:  
; APPLICANT: Cochran Ph.D., Mark D  
; APPLICANT: Junker M.S., David E  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07820,154A  
; APPLICATION NUMBER: US/07820,154A  
; FILING DATE: 19920113  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)977-9550  
; TELEFAX: (212)664-0525  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-820-154A-12

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 X 1

Db 1 V 1

RESULT 2

US-07-791-213D-24

Sequence 24, Application US/07791213D  
Patent No. 5409895  
GENERAL INFORMATION:  
APPLICANT: MORISHITA, Hideaki  
APPLICANT: KANAMORI, Toshinori  
APPLICANT: NOBUHARA, Masahiro  
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
TITLE OF INVENTION: TREATING USING THE SAME  
NUMBER OF SEQUENCES: 108  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07791,213D  
FILING DATE: 13-NOV-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-306745  
FILING DATE: 13-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 029650-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 C 1

RESULT 4  
US-08-174-365A-57  
Sequence 57, Application US/08174365A  
Patent No. 5478809  
GENERAL INFORMATION:  
APPLICANT: Seichi TANIDA et al.  
TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/174,365A  
FILING DATE: December 28, 1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 57:

Qy 1 X 1  
Db 1 C 1

RESULT 3  
US-07-791-213D-40  
Sequence 40, Application US/07791213D  
Patent No. 5409895  
GENERAL INFORMATION:  
APPLICANT: MORISHITA, Hideaki  
APPLICANT: KANAMORI, Toshinori  
APPLICANT: NOBUHARA, Masahiro  
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
TITLE OF INVENTION: TREATING USING THE SAME  
NUMBER OF SEQUENCES: 108  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 1 amino acid

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: modified site

; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: /note = "Xaa is modified amino acid as

; OTHER INFORMATION: described in specification"

US-08-174-365A-57

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1

DB 1 X 1

## RESULT 5

US-07-789-913-23

; Sequence 23, Application US/07789913

; Patent No. 5559095

; GENERAL INFORMATION:

; APPLICANT: Miljanich, George P.

; APPLICANT: Bowersox, Stephen S.

; APPLICANT: Fox, James A.

; APPLICANT: Valentino, Karen L.

; APPLICANT: Bitner, Robert S.

; APPLICANT: Yamashiro, Donald H.

; TITLE OF INVENTION: Delayed Treatment Method of Reducing

; TITLE OF INVENTION: Ischemia-Related Neuronal Damage

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Law Offices of Peter Dehlinger

; STREET: 350 Cambridge Avenue, Suite 300

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/789,913

; FILING DATE: 19911112

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/561,766

; FILING DATE: 02-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/440,094

; FILING DATE: 22-NOV-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Stratford, Carol A.

; REGISTRATION NUMBER: 34,444

; REFERENCE/DOCKET NUMBER: 5865-0005.30

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: both

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: peptide fragment used in the claims

US-07-789-913-23

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 C 1

## RESULT 6

US-07-789-913-25

; Sequence 25, Application US/07789913

; Patent No. 5559095

; GENERAL INFORMATION:

; APPLICANT: Miljanich, George P.

; APPLICANT: Bowersox, Stephen S.

; APPLICANT: Fox, James A.

; APPLICANT: Valentino, Karen L.

; APPLICANT: Bitner, Robert S.

; APPLICANT: Yamashiro, Donald H.

; TITLE OF INVENTION: Delayed Treatment Method of Reducing

; TITLE OF INVENTION: Ischemia-Related Neuronal Damage

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Law Offices of Peter Dehlinger

; STREET: 350 Cambridge Avenue, Suite 300

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/789,913

; FILING DATE: 19911112

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/561,766

; FILING DATE: 02-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/440,094

; FILING DATE: 22-NOV-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Stratford, Carol A.

; REGISTRATION NUMBER: 34,444

; REFERENCE/DOCKET NUMBER: 5865-0005.30

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: both

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: peptide fragment used in the claims

US-07-789-913-25

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 R 1

RESULT 7

US-08-049-794-23

Sequence 23, Application US/08049794

Patent No. 5587454

GENERAL INFORMATION:

APPLICANT: JUSTICE, ALAN

APPLICANT: SINGH, TEJINDER

APPLICANT: GOHIL, KISHOR C

APPLICANT: VALENTINO, KAREN L

APPLICANT: MILJANICH, GEORGE P

TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Peter Dehlinger

STREET: 350 Cambridge Avenue, Suite 300

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/049,794

FILING DATE: 19930415

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/814,759

FILING DATE: 30-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A.

REGISTRATION NUMBER: 34,444

REFERENCE/DOCKET NUMBER: 5865-0009.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0960

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE

INDIVIDUAL ISOLATE: 32

US-08-049-794-23

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 C 1

RESULT 8

US-08-049-794-25

Sequence 25, Application US/08049794

Patent No. 5587454

GENERAL INFORMATION:

APPLICANT: JUSTICE, ALAN

APPLICANT: SINGH, TEJINDER

APPLICANT: GOHIL, KISHOR C

APPLICANT: VALENTINO, KAREN L

TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Peter Dehlinger

STREET: 350 Cambridge Avenue, Suite 300

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/049,794

FILING DATE: 19930415

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/814,759

FILING DATE: 30-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A.

REGISTRATION NUMBER: 34,444

REFERENCE/DOCKET NUMBER: 5865-0009.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0960

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE

INDIVIDUAL ISOLATE: 32

US-08-049-794-23

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 C 1

RESULT 9

US-08-433-037-12

Sequence 12, Application US/08433037

Patent No. 5707828

GENERAL INFORMATION:

APPLICANT: Sreekrishna, Kotikanyadan

APPLICANT: Barr, Kathryn A.

APPLICANT: Brierley, Russell A.

APPLICANT: Thill, Gregory P.

APPLICANT: Tschoop, Juerg F.

TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: U.S.A.

ZIP: 11530-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

US-08-049-794-25

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 R 1

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 91082  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-12

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 L 1

RESULT 10  
US-08-448-606-4  
Sequence 4, Application US/08448606  
Patent No. 5721114  
GENERAL INFORMATION:  
APPLICANT: Abrahams n, Lars  
APPLICANT: Holmgren, Erik  
APPLICANT: Kalder n, Christina  
APPLICANT: Lake, Mats  
APPLICANT: Mikaelsson, sa  
APPLICANT: Sejlitz, Torsten  
TITLE OF INVENTION: Expression System For Producing  
TITLE OF INVENTION: Apolipoprotein AI-M  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pollock, Vande Sande & Priddy  
STREET: 1990 M Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,606  
FILING DATE: 25-AUG-1995  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE93/01061  
FILING DATE: 09-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9203753-0  
FILING DATE: 11-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Amernick, Burton A.  
REGISTRATION NUMBER: 24,852  
REFERENCE/DOCKET NUMBER: 0151/00121  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)331-7111  
TELEFAX: (202) 293-6229  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acid  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-448-606-4

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 Q 1

RESULT 11  
US-07-869-933-16  
Sequence 16, Application US/07869933  
Patent No. 5770396  
GENERAL INFORMATION:  
APPLICANT: KINET, Jean-Pierre  
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF  
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR  
TITLE OF INVENTION: IMMUNOGLOBULIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/869,933  
FILING DATE: 19920416  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40399/154 NIHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-869-933-16

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 M 1

RESULT 12  
US-08-293-150A-24  
Sequence 24, Application US/08293150A

Patent No. 5792629  
GENERAL INFORMATION:  
APPLICANT: MORISHITA, Hideaki  
APPLICANT: KANAMORI, Toshinori  
APPLICANT: NOBUHARA, Masahiro  
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
TITLE OF INVENTION: TREATING USING THE SAME  
NUMBER OF SEQUENCES: 110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/293,150A  
FILING DATE: 19-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/791,213  
FILING DATE: 13-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-306745  
FILING DATE: 13-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 029650-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-293-150A-24

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 C 1

RESULT 13  
US-08-293-150A-40  
Sequence 40, Application US/08293150A  
Patent No. 5792629  
GENERAL INFORMATION:  
APPLICANT: MORISHITA, Hideaki  
APPLICANT: KANAMORI, Toshinori  
APPLICANT: NOBUHARA, Masahiro  
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
TITLE OF INVENTION: TREATING USING THE SAME  
NUMBER OF SEQUENCES: 110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria

STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/293,150A  
FILING DATE: 19-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/791,213  
FILING DATE: 13-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-306745  
FILING DATE: 13-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 029650-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-293-150A-40

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 C 1

RESULT 14  
US-08-496-847-23  
Sequence 23, Application US/08496847  
Patent No. 5795864  
GENERAL INFORMATION:  
APPLICANT: Ametutz, Gary A.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Gobil, Kishorchandra  
APPLICANT: Adriaenssens, Peter I.  
APPLICANT: Kristipati, Ramasharma  
TITLE OF INVENTION: METHODS AND  
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94306-1546  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,847  
FILING DATE: 27-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A



REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0009.31  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT  
US-08-496-847-23

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 C 1

## RESULT 15

US-08-496-847-25  
Sequence 25, Application US/08496847  
Patent No. 5795864

## GENERAL INFORMATION:

APPLICANT: Amstutz, Gary A.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Gohil, Kishorchandra  
APPLICANT: Adriaenssens, Peter I.  
APPLICANT: Kristipati, Ramasharma

TITLE OF INVENTION: METHODS AND

FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94306-1546

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/496,847

FILING DATE: 27-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A

REGISTRATION NUMBER: 34,444

REFERENCE/DOCKET NUMBER: 5865-0009.31

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT  
US-08-496-847-25

Query Match

0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 R 1

Search completed: December 8, 2003, 09:20:29  
Job time : 1.85859 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 3.55556 Seconds  
(without alignments)  
209.232 Million cell updates/sec

Title: US-09-498-556C-85  
Perfect score: 4  
Sequence: 1 XXXX 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	9 US-09-909-348-4	Sequence 4, Appli
2	0	0.0	1	10 US-09-982-172-3	Sequence 3, Appli
3	0	0.0	1	10 US-09-982-172-4	Sequence 4, Appli
4	0	0.0	1	10 US-09-982-172-9	Sequence 9, Appli
5	0	0.0	1	10 US-09-982-172-11	Sequence 11, Appli
6	0	0.0	1	10 US-09-982-172-19	Sequence 19, Appli
7	0	0.0	1	10 US-09-982-172-31	Sequence 31, Appli
8	0	0.0	1	10 US-09-982-172-35	Sequence 35, Appli
9	0	0.0	1	10 US-09-982-172-37	Sequence 37, Appli
10	0	0.0	1	10 US-09-982-172-46	Sequence 46, Appli
11	0	0.0	1	10 US-09-982-172-69	Sequence 69, Appli
12	0	0.0	1	10 US-09-982-172-80	Sequence 80, Appli
13	0	0.0	1	10 US-09-982-172-81	Sequence 81, Appli
14	0	0.0	1	10 US-09-982-172-83	Sequence 83, Appli
15	0	0.0	1	10 US-09-982-172-86	Sequence 86, Appli

16	0	0.0	1	10 US-09-982-172-93	Sequence 93, Appli
17	0	0.0	1	10 US-09-982-172-95	Sequence 95, Appli
18	0	0.0	1	10 US-09-982-172-106	Sequence 106, Appli
19	0	0.0	1	10 US-09-982-172-112	Sequence 112, Appli
20	0	0.0	1	10 US-09-982-172-120	Sequence 120, Appli
21	0	0.0	1	10 US-09-982-172-126	Sequence 126, Appli
22	0	0.0	1	10 US-09-982-172-148	Sequence 148, Appli
23	0	0.0	1	10 US-09-982-172-149	Sequence 149, Appli
24	0	0.0	1	10 US-09-982-172-155	Sequence 155, Appli
25	0	0.0	1	10 US-09-982-172-160	Sequence 160, Appli
26	0	0.0	1	10 US-09-982-172-172	Sequence 172, Appli
27	0	0.0	1	10 US-09-982-172-173	Sequence 173, Appli
28	0	0.0	1	10 US-09-982-172-175	Sequence 175, Appli
29	0	0.0	1	10 US-09-982-172-189	Sequence 189, Appli
30	0	0.0	1	10 US-09-982-172-190	Sequence 190, Appli
31	0	0.0	1	10 US-09-982-172-191	Sequence 191, Appli
32	0	0.0	1	10 US-09-982-172-195	Sequence 195, Appli
33	0	0.0	1	10 US-09-982-172-200	Sequence 200, Appli
34	0	0.0	1	10 US-09-982-172-211	Sequence 211, Appli
35	0	0.0	1	11 US-09-809-391-395	Sequence 395, Appli
36	0	0.0	1	11 US-09-809-391-611	Sequence 611, Appli
37	0	0.0	1	12 US-10-032-221B-19	Sequence 19, Appli
38	0	0.0	1	12 US-10-299-619-4	Sequence 4, Appli
39	0	0.0	1	12 US-10-061-607A-28	Sequence 28, Appli
40	0	0.0	1	12 US-10-061-607A-29	Sequence 29, Appli
41	0	0.0	1	12 US-10-061-607A-35	Sequence 35, Appli
42	0	0.0	1	12 US-09-882-171-395	Sequence 395, Appli
43	0	0.0	1	12 US-09-882-171-611	Sequence 611, Appli
44	0	0.0	1	12 US-10-104-307-7	Sequence 7, Appli
45	0	0.0	1	12 US-10-104-307-8	Sequence 8, Appli

## ALIGNMENTS

### RESULT 1

US-09-909-348-4  
; Sequence 4, Application US/0909348  
; Patent No. US20020042373A1  
; GENERAL INFORMATION:  
; APPLICANT: Carney, Darrell H.  
; APPLICANT: Crowther, Roger S.  
; APPLICANT: Stierberg, Janet  
; APPLICANT: Bergmann, John  
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists  
; FILE REFERENCE: Of The No. US20020042373A1-Proteolytically Activated Thrombin R  
; CURRENT APPLICATION NUMBER: US/09/909,348  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: US 60/219,800  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide fragment of Thrombin  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(14)  
; OTHER INFORMATION: Xaa at position six is Glu or Gln  
; OTHER INFORMATION: Xaa at position thirteen is Phe, Met, Leu, His or  
US-09-909-348-4

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 V 1

```
RESULT 2
US-09-982-172-3
; Sequence 3, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-3

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 3
US-09-982-172-4
; Sequence 4, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-4

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 4
US-09-982-172-9
; Sequence 9, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
```

```
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-9

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 5
US-09-982-172-11
; Sequence 11, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-11

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 6
US-09-982-172-19
; Sequence 19, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-19
```

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 X 1  
DB 1 K 1  
  
RESULT 7  
US-09-982-172-31  
; Sequence 31, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL  
; FILE REFERENCE: UTILIZING EACH  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-31

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 X 1  
DB 1 K 1

RESULT 8  
US-09-982-172-35  
; Sequence 35, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL  
; FILE REFERENCE: UTILIZING EACH  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-35

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 X 1  
DB 1 R 1

RESULT 9

US-09-982-172-37  
; Sequence 37, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL  
; FILE REFERENCE: UTILIZING EACH  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 37  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-37

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 X 1  
DB 1 K 1

RESULT 10  
US-09-982-172-46  
; Sequence 46, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL  
; FILE REFERENCE: UTILIZING EACH  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-46

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 X 1  
DB 1 K 1

RESULT 11  
US-09-982-172-69  
; Sequence 69, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL  
; FILE REFERENCE: UTILIZING EACH  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19

; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 69  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-69

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

RESULT 12  
US-09-982-172-80  
; Sequence 80, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 80  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-80

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

RESULT 13  
US-09-982-172-81  
; Sequence 81, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 81  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-81

Query Match 0.0%; Score 0; DB 10; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

US-09-982-172-83  
; Sequence 83, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 83  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-83

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

US-09-982-172-86  
; Sequence 86, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-86

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

US-09-982-172-86  
; Sequence 86, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-86

Query Match 0.0%; Score 0; DB 10; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

US-09-982-172-83  
; Sequence 83, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 83  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-83

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

US-09-982-172-86  
; Sequence 86, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-86

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

US-09-982-172-86  
; Sequence 86, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; TITLE OF INVENTION: UTILIZING EACH  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-86

Query Match 0.0%; Score 0; DB 10; Length 1;

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	125	25.7	561	2	T27318	hypothetical prote
2	125	25.7	626	2	T27319	hypothetical prote
3	113	23.3	490	2	T32003	hypothetical prote
4	112	23.0	249	2	T24604	hypothetical prote
5	109.5	22.5	1642	2	T19130	hypothetical prote
6	105.5	21.7	145	2	T15608	hypothetical prote
7	100	20.6	5376	2	T42315	zonadhesin - mouse
8	99	20.4	1036	2	T17405	scavenger receptor
9	98.5	20.3	135	2	T15610	hypothetical prote
10	98	20.2	137	2	T15609	hypothetical prote
11	96.5	19.9	869	1	JC4858	VLDR receptor prec
12	95.5	19.7	166	2	H89054	protein B0238.12 f
13	94.5	19.4	98	2	C89046	protein C1038.4 [i
14	94.5	19.4	4660	2	T42737	Sp330 protein prec
15	92.5	19.0	63	2	S07127	chymotrypsin/elast
16	92.5	19.0	63	2	S08572	chymotrypsin/elast
17	92.5	19.0	4753	1	A47437	LDL-receptor-relat
18	90.5	18.6	195	2	T28803	hypothetical prote
19	87.5	18.0	802	2	T24293	hypothetical prote
20	87.5	18.0	949	2	T24294	hypothetical prote
21	85.5	17.6	863	1	S51789	VLDR receptor prec
22	85	17.5	2155	2	T30197	alpha tectorin - m
23	84.5	17.4	13288	2	T03099	mucin, submaxillar
24	83	17.1	209	2	T02394	hypothetical prote
25	82.5	17.0	873	1	A49729	VLDR receptor prec
26	82.5	17.0	873	1	QR8BVD	VLDR receptor prec
27	82	16.9	62	2	S35098	trypsin inhibitor
28	82	16.9	267	2	T32007	hypothetical prote
29	81.5	16.8	1810	1	A32320	tenascin precursor





```

C;Accession: T15610
R;Bradshaw, H.
A;Description: The sequence of C. elegans cosmid C25E10.
A;Reference number: Z18376
A;Accession: T15610
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-135 <BRA>
A;Cross-references: EMBL:U50311; NID:gl1226295; PID:gl1226304; PIDN:AAA92314.1; CESP:
C;Genetics:
A;Gene: CESP:C25E10.9
A;Introns: 1/3; 18/1; 41/3

Query Match      20.3%; Score 98.5; DB 2; Length 135;
Best Local Similarity 32.0%; Pred. No. 0.022;
Matches 24; Conservative 11; Mismatches 19; Indels 21; Gaps 5;

Qy 1 KATWQGENEKYDCSGKCKKCKYDGVVEEDDEEPN-----VPCLVRVCHQDCVCEEG 55
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 75 KETTTCPENETFFGCGT-ACEATC-----EKPNTVCTKQCIWVCQ-----CSKG 119
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 56 FYRNKODKCVSAEDC 70
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 FVRH-GLRCIDKKDC 133
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 10
T15609
hypothetical protein C25E10.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15609
R;Bradshaw, H.
A;Description: The sequence of C. elegans cosmid C25E10.
A;Reference number: Z18376.
A;Accession: T15609
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-137 <BRA>
A;Cross-references: EMBL:U50311; NID:gl1226295; PID:gl1226303; PIDN:AAA92313.1; CESP:
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:C25E10.8
A;Introns: 1/3; 19/1; 42/3

Query Match      20.2%; Score 98; DB 2; Length 137;
Best Local Similarity 31.8%; Pred. No. 0.025;
Matches 28; Conservative 9; Mismatches 33; Indels 18; Gaps 5;

Qy 2 ATWQGENEKYDCSGKCKKCKYDGVVEEDDEEPNVPCLVRVCHQD-CVCEGFRNK 60
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 17 ANKTCGANEEWVAC-HDHCTQCGY-----TPKV-CLSAQCIENACDCKKGFVRNS 65
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 61 DDKCVSAEDCEL-----DNMDFIYPGT 82
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 66 LGKVDVISTCKETSKCPNETFFRCGT 93
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 11
JC4858
VLDL receptor precursor - African clawed frog
N;Alternate names: very low density lipoprotein receptor; vitellogenin receptor
C;Species: Xenopus laevis (African clawed frog)
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 16-Jun-2000
C;Accession: JC4858
R;Okabayashi, K.; Shoji, H.; Nakamura, T.; Hashimoto, O.; Asashima, M.; Sugino, H.
Biochem. Biophys. Res. Commun. 224, 406-413, 1996
A;Title: cDNA cloning and expression of the Xenopus laevis vitellogenin receptor.
A;Reference number: JC4858; MUID:96295501; PMID:8702402
A;Accession: JC4858

```

A:Molecule type: mRNA  
A:Residues: 1-869 <OKA>  
A:Cross-references: GB:AB006906; NID:g2366772; PIDN:BAA22145.1; PID:g3366773  
C:Comment: This receptor mediates incorporation of vitellogenin into oocytes.  
C:Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;  
C:Keywords: duplication; fatty acid metabolism; glycoprotein; receptor; transmembrane p  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-869/Product: VLDL receptor #status predicted <MAT>  
F:27-793/Domain: extracellular #status predicted <EXT>  
F:32-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:71-107/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F:112-148/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F:153-187/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F:132-228/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F:238-272/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F:277-311/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F:317-354/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F:359-393/Domain: EGF homology <EG1>  
F:399-433/Domain: EGF homology <EG2>  
F:440-479/Domain: LDL receptor WYTD-containing repeat homology <YW1>  
F:480-525/Domain: LDL receptor WYTD-containing repeat homology <YW2>  
F:526-568/Domain: LDL receptor WYTD-containing repeat homology <YW3>  
F:569-612/Domain: LDL receptor WYTD-containing repeat homology <YW4>  
F:613-655/Domain: LDL receptor WYTD-containing repeat homology <YW5>  
F:656-698/Domain: LDL receptor WYTD-containing repeat homology <YW6>  
F:707-749/Domain: EGF homology <EG3>  
F:794-815/Domain: transmembrane #status predicted <TM>  
F:816-869/Domain: intracellular #status predicted <CYT>  
F:830-834/Region: coated-pit mediated internalization signal  
F:150,201,777,786/Binding site: carbohydrate (Agn) (covalent) #status predicted  
F:359-370,366-379,381-393,399-409,405-418,420-433,707-720,716-735,737-749/Disulfide bond

Query Match 19.9%; Score 96.5; DB 1; Length 869;  
Best Local Similarity 35.4%; Pred. No. 0.16;  
Matches 28; Conservative 8; Mismatches 38; Indels 5; Gaps 4;

Qy 2 ATMCGGENEKYDCSGKCEC-DKKCKYDGVVEEDDEPNVPLVVRVCHOD-CVCEGFGYRN 59  
Db 234 APQRCANEM--PCGSECTHKRWKCDGADCKRDSDEINCPSTCTCPDQFKCEDGNCIH 291

Qy 60 KDKCVSAEDCELDNMDFI 78  
Db 292 GSRQCDGVRDC-LDGTDEI 309

RESULT 12  
H89044  
protein B0238.12 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: H89044  
R:anonymous, The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: H89044  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-166 <STO>  
A:Cross-references: GB:chr\_V; PIDN:AAB65990.1; PID:g2315490; GSPDB:GN00023; CESP:B0238.1  
C:Genetics:  
A:Map position: 5

Query Match 19.7%; Score 95.5; DB 2; Length 166;  
Best Local Similarity 31.4%; Pred. No. 0.049;  
Matches 22; Conservative 9; Mismatches 28; Indels 11; Gaps 4;

Qy 1 KATMCGGENEKYDCSGKCEC-DKKCKYDGVVEEDDEPNVPLVVRVCHODCVCEGFGYRNK 60  
Db 108 QSPQRCGRNETFRTGCS-SCFPSC-----TTPRQACTMQCIVNVQV-----CSEGFVRGP 157

Qy 61 DDKCVSAEDC 70  
Db 158 SG-CVRQRDC 166

RESULT 13  
C89046  
protein C10G8.4 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: C89046  
R:anonymous, The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: C89046  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <STO>  
A:Cross-references: GB:chr\_V; PIDN:AAB09171.1; PID:g1572829; GSPDB:GN00023; CESP:C10G8.4  
C:Genetics:  
A:Gene: C10G8.4  
A:Map position: 5

Query Match 19.4%; Score 94.5; DB 2; Length 98;  
Best Local Similarity 28.8%; Pred. No. 0.039;  
Matches 23; Conservative 13; Mismatches 17; Indels 27; Gaps 6;

Qy 5 QCGENKDYDCSGKCEC-DKKCKYDGVVEEDDEPN-----VPLVVRVCHODCVCEGFGYRN 59  
Db 39 RCPSENEEFKSCGT-ACEPTC-----QAPNPQVCTIQCLINVCQ-----CSQGFVRG 83

Qy 60 KDKCVSAEDCELDNMDFIY 79  
Db 84 PNG-CVPPQDC-----FVY 96

RESULT 14  
T42737  
gp330 protein precursor - rat  
N:Alternate names: megalin  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C:Accession: T42737  
R:Saito, A.; Pietromonaco, S.; Loo, A.K.C.; Farquhar, M.G.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994  
A:Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of ti  
A:Reference number: A58173; MUID:95024033; PMID:7937880  
A:Accession: T42737  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-4660 <SAI>  
A:Cross-references: EMBL:L34049; NID:g561852; PID:g561853; PIDN:AAA51369.1  
A:Experimental source: strain Sprague-Dawley; kidney  
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-4660/Product: gp330 protein #status predicted <MAT>

Query Match 19.4%; Score 94.5; DB 2; Length 4660;  
Best Local Similarity 30.0%; Pred. No. 0.97;  
Matches 30; Conservative 11; Mismatches 24; Indels 35; Gaps 8;

Qy 6 CGENKDYDC-----GSKEC---DKKCKYDGVVEE---EDDEPNV-----PCLVRY--- 45  
Db 256 CESNQSHHRCYPREWACPGSGRCISIDKVC--DGVPCPEGDENNVTSGRTCGMGVCSV 313

Qy 46 -----CHQ-----DCVCEGFGYRNKDDK---CVSAEDCEL 72  
Db 314 LNCYQCHQTFPGEGCFPCPGHINSNSRSTCTDFDDCOI 353



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 / Search time 5.45455 Seconds  
(without alignments)  
116.399 Million cell updates/sec

Title: US-09-498-556C-85  
Perfect score: 4  
Sequence: 1 XXXX 4

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database: A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	20 AAY46652	Immunogenic peptide
2	0	0.0	1	22 ABB56870	Human SNP related
3	0	0.0	1	22 ABB66809	Drosophila melanog
4	0	0.0	1	22 ABB66810	Drosophila melanog
5	0	0.0	1	22 ABB66811	Novel human diagno
6	0	0.0	1	22 ABB66812	Human peptide #918
7	0	0.0	1	22 ABB66813	Human peptide #110
8	0	0.0	1	22 ABB66814	Human peptide #124
9	0	0.0	1	22 ABB66815	Human peptide #162

10	0	0.0	1	22 AAM98447	Human peptide #172
11	0	0.0	1	22 AAM53218	Human nonconservat
12	0	0.0	1	22 AAM53219	Human nonconservat
13	0	0.0	1	22 AAM53290	Human nonconservat
14	0	0.0	1	22 AAM53291	Human nonconservat
15	0	0.0	1	22 AAM53328	Human nonconservat
16	0	0.0	1	22 AAM53329	Human nonconservat
17	0	0.0	1	22 AAM53330	H11 binding site c
18	0	0.0	1	22 AAG99966	ERA binding domain
19	0	0.0	1	22 AAG99983	ERA binding domain
20	0	0.0	1	22 AAG99987	ERA binding domain
21	0	0.0	1	22 AAG99988	ERA binding domain
22	0	0.0	1	22 AAM00010	ERA binding domain
23	0	0.0	1	22 AAM00011	ERA binding domain
24	0	0.0	1	22 AAM00013	ERA binding domain
25	0	0.0	1	22 AAM00016	ERA binding domain
26	0	0.0	1	22 AAG98026	Human SNP associat
27	0	0.0	1	22 AAG98134	Thyrotropin releas
28	0	0.0	1	22 AAB91029	Endothelins and re
29	0	0.0	1	22 AAB91546	Opioid peptide SEQ
30	0	0.0	1	22 AAB91665	Opioid peptide SEQ
31	0	0.0	1	22 AAB91739	Apoptosis related
32	0	0.0	1	22 AAB91892	Polypeptide SEQ ID
33	0	0.0	1	22 AAB92150	Miscellaneous pept
34	0	0.0	1	22 AAB92392	Helicobacter pylor
35	0	0.0	1	23 ABUS1078	Helicobacter pylor
36	0	0.0	1	23 ABUS1919	Helicobacter pylor
37	0	0.0	1	23 ABUS2012	Helicobacter pylor
38	0	0.0	1	23 ABUS2118	Human novel secret
39	0	0.0	1	23 ABG95255	Human novel secret
40	0	0.0	1	23 ABG95471	Human P-glycoprote
41	0	0.0	1	23 ABG93483	Human P-glycoprote
42	0	0.0	1	23 ABG93484	Human P-glycoprote
43	0	0.0	1	23 ABG93489	Human P-glycoprote
44	0	0.0	1	23 ABG93491	Human P-glycoprote
45	0	0.0	1	23 ABG93499	Human P-glycoprote

## ALIGNMENTS

RESULT 1  
AAY46652  
ID AAY46652 standard; peptide; 1 AA.  
XX AAY46652;  
AC AAY46652;  
XX  
DT 01-DEC-1999 (first entry)  
XX  
DE Immunogenic peptide having a human leukocyte antigen binding motif #1263.  
XX  
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
KW Immune response; T cell activation; major histocompatibility complex;  
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
KW vaccine; immunisation.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO9945954-A1.  
XX  
PD 16-SEP-1999.  
XX  
PF 13-MAR-1998; 98WO-US05039.  
XX  
PR 13-MAR-1998; 98WO-US05039.  
XX  
PA (EPIM-) EPIMUNE INC.  
XX  
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
XX WPI; 1999-551214/46.  
DR



CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 H 1

RESULT 4  
ABB66810  
ID ABB66810 standard; Protein; 1 AA.

XX AC ABB66810;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 27222.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical.

XX OS Drosophila melanogaster.

XX FN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR N-PSDB; ABL10913.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -

XX Disclosure; SEQ ID NO 27222; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
sequences (ABL01840-ABL16175) and the encoded proteins  
(ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 H 1

RESULT 5  
ABG02941  
ID ABG02941 standard; Protein; 1 AA.

XX AC ABG02941;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #2932.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX DR N-PSDB; AAS67128.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -

XX Claim 20; SEQ ID NO 33300; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags  
for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. ABG00010-ABG30377 represent novel human  
diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

```

Db      1 M 1

RESULT 6
AAM97643
ID AAM97643 standard; Peptide; 1 AA.
XX
AC AAM97643;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #918 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Disclosure; Page 3869; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 7
AAM97834
ID AAM97834 standard; Peptide; 1 AA.
XX
AC AAM97834;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #1109 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Disclosure; Page 3911; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 E 1

RESULT 8
AAM97974

```





DE Human peptide #1722 encoded by a SNP oligonucleotide.

KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;

XX neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;

KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;

KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;

KW complement related protein; cytochrome; kinesin; cytokine; interferon;

KW interleukin; G-protein coupled receptor; thioesterase; inflammation;

KW multifactorial disease; autoimmune disease; infection;

KW nervous system disease.

XX Homo sapiens.

OS WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2001-465210/50.

DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

XX oncogenes and histones, useful for diagnosing and treating, e.g.

PT cancer, autoimmune diseases and infections

XX Disclosure; Page 4045; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)

CC encoding polymorphic variants of proteins related to amyloid, amyloid

CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,

CC polymerase, oncogenes, histones, kinases, colony stimulating factors,

CC complement related proteins, cytochromes, kinesins, cytokines,

CC interferons, interleukins, G-protein coupled receptors and thioesterases.

CC The present sequence is a peptide encoded by one such oligonucleotide.

CC The oligonucleotides and the peptides encoded by them may be used in the

CC prevention, diagnosis and treatment of diseases associated with

CC inappropriate expression of the proteins listed above. Disorders that may

CC be prevented, diagnosed and/or treated include multifactorial diseases

CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid

CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus

CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,

CC brain, breast, colon and kidney, leukaemia), diseases of the nervous

CC system and an infection of pathogenic organisms.

XX Sequence 1 AA;

Qy Query Match 0.0%; Score 0; DB 22; Length 1;

Db Best Local Similarity 0.0%; Pred.No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 L 1

RESULT 11

AAM53218

ID AAM53218 standard; Peptide; 1 AA.

XX AAM53218;

AC AAM53218;

XX 09-NOV-2001 (first entry)

DT Human nonconservative amino acid changing SNP related peptide SEQ:6913.

DE Human; single nucleotide polymorphism; SNP; genome; gene therapy;

XX

KW

protein therapy; vaccine; probe; diagnostic assay; detection;

quantitation; restorative therapy; polymorphic.

Homo sapiens.

WO200140521-A2.

07-JUN-2001.

30-NOV-2000; 2000WO-US32758.

30-NOV-1999; 99US-0168138.

29-NOV-2000; 2000US-0726173.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach M;

WPI; 2001-356160/37.

Polymorphic nucleic acid sequences, useful in genetic testing and

therapy -

Claim 29; Page 2619; 2653pp; English.

AAT73060 to AAT79867 represent isolated human polymorphic polynucleotide

sequences (I), which contain single nucleotide polymorphisms (SNPs).

AAM53114 to AAM53329 represent peptides related to human polymorphic

polynucleotide sequences. The sequences can be used in gene and protein

therapy, and in vaccine production. (I) and the polypeptides encoded by

them may be used in the prevention, diagnosis and treatment of diseases

associated with inappropriate expression of polymorphic polypeptides.

For example, (I) may be used to treat disorders by rectifying mutations

or deletions in a patient's genome that affect the activity of

polypeptides by expressing inactive proteins or to supplement the

patients own production of polypeptide. Additionally, (I) and its

complementary sequences may also be used as DNA probes in diagnostic

assays to detect and quantitate the presence of similar nucleic acids

in samples, and therefore which patients may be in need of restorative

therapy. The polypeptides encoded by (I) may be used as antigens in the

production of antibodies specific for polymorphic polypeptides. The

antibodies may also be used to down regulate expression and activity.

The antibodies may also be used as diagnostic agents for detecting the

presence of polymorphic polypeptides in samples.

Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;

Best Local Similarity 0.0%; Pred.No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 H 1

RESULT 12

AAM53219

ID AAM53219 standard; Peptide; 1 AA.

XX AAM53219;

AC AAM53219;

XX 09-NOV-2001 (first entry)

DT Human nonconservative amino acid changing SNP related peptide SEQ:6914.

DE Human; single nucleotide polymorphism; SNP; genome; gene therapy;

KW protein therapy; vaccine; probe; diagnostic assay; detection;

KW quantitation; restorative therapy; polymorphic.

XX Homo sapiens.

OS WO200140521-A2.

PN

```
XX PD 07-JUN-2001.
XX XX
XX PF 30-NOV-2000; 2000WO-US32758.
XX PI
XX PR 30-NOV-1999; 99US-0168138.
XX PR 29-NOV-2000; 2000US-0726173.
XX XX
XX PA (CURA-) CURAGEN CORP.
XX XX Shimkets RA, Leach M;
XX PI WPI; 2001-356160/37.
XX DR
XX XX
XX PF Polymorphic nucleic acid sequences, useful in genetic testing and
XX PT therapy -
XX XX
XX PS Claim 29; Page 2619; 2653pp; English.
XX CC
XX CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
XX CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
XX CC AAM53114 to AAM53329 represent peptides related to human polymorphic
XX CC polynucleotide sequences. The sequences can be used in gene and protein
XX CC therapy, and in vaccine production. (I) and the polypeptides encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate expression of polymorphic polypeptides.
XX CC For example, (I) may be used to treat disorders by rectifying mutations
XX CC or deletions in a patient's genome that affect the activity of
XX CC polypeptides by expressing inactive proteins or to supplement the
XX CC patients own production of polypeptide. Additionally, (I) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays to detect and quantitate the presence of similar nucleic acids
XX CC in samples, and therefore which patients may be in need of restorative
XX CC therapy. The polypeptides encoded by (I) may be used as antigens in the
XX CC production of antibodies specific for polymorphic polypeptides. The
XX CC antibodies may also be used to down regulate expression and activity.
XX CC The antibodies may also be used as diagnostic agents for detecting the
XX CC presence of polymorphic polypeptides in samples.
XX SQ Sequence 1 AA;
XX
XX Query Match 0.0%; Score 0; DB 22; Length 1;
XX Best Local Similarity 0.0%; Pred. No. 0;
XX Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 X 1
XX Db 1 H 1
XX
XX RESULT 13
XX AAM53290
XX ID AAM53290 standard; Peptide; 1 AA.
XX AC AAM53290;
XX XX
XX DT 09-NOV-2001 (first entry)
XX DE Human nonconservative amino acid changing SNP related peptide SEQ:6985.
XX XX
XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
XX KW protein therapy; vaccine; probe; diagnostic assay; detection;
XX KW quantitation; restorative therapy; polymorphic.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200140521-A2.
XX XX
XX PD 07-JUN-2001.
XX PF 30-NOV-2000; 2000WO-US32758.
XX PR 30-NOV-1999; 99US-0168138.
XX XX
XX PA (CURA-) CURAGEN CORP.
XX XX Shimkets RA, Leach M;
XX PI WPI; 2001-356160/37.
XX DR
```

```
PR 29-NOV-2000; 2000US-0726173.
XX (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX XX WPI; 2001-356160/37.
XX XX
XX PT Polymorphic nucleic acid sequences, useful in genetic testing and
XX PT therapy -
XX XX
XX PS Claim 29; Page 2641; 2653pp; English.
XX CC
XX CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
XX CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
XX CC AAM53114 to AAM53329 represent peptides related to human polymorphic
XX CC polynucleotide sequences. The sequences can be used in gene and protein
XX CC therapy, and in vaccine production. (I) and the polypeptides encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate expression of polymorphic polypeptides.
XX CC For example, (I) may be used to treat disorders by rectifying mutations
XX CC or deletions in a patient's genome that affect the activity of
XX CC polypeptides by expressing inactive proteins or to supplement the
XX CC patients own production of polypeptide. Additionally, (I) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays to detect and quantitate the presence of similar nucleic acids
XX CC in samples, and therefore which patients may be in need of restorative
XX CC therapy. The polypeptides encoded by (I) may be used as antigens in the
XX CC production of antibodies specific for polymorphic polypeptides. The
XX CC antibodies may also be used to down regulate expression and activity.
XX CC The antibodies may also be used as diagnostic agents for detecting the
XX CC presence of polymorphic polypeptides in samples.
XX SQ Sequence 1 AA;
XX
XX Query Match 0.0%; Score 0; DB 22; Length 1;
XX Best Local Similarity 0.0%; Pred. No. 0;
XX Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 X 1
XX Db 1 L 1
XX
XX RESULT 14
XX AAM53291
XX ID AAM53291 standard; Peptide; 1 AA.
XX AC AAM53291;
XX XX
XX DT 09-NOV-2001 (first entry)
XX DE Human nonconservative amino acid changing SNP related peptide SEQ:6986.
XX XX
XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
XX KW protein therapy; vaccine; probe; diagnostic assay; detection;
XX KW quantitation; restorative therapy; polymorphic.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200140521-A2.
XX XX
XX PD 07-JUN-2001.
XX PF 30-NOV-2000; 2000WO-US32758.
XX PR 30-NOV-1999; 99US-0168138.
XX PR 29-NOV-2000; 2000US-0726173.
XX XX
XX PA (CURA-) CURAGEN CORP.
XX XX Shimkets RA, Leach M;
XX PI WPI; 2001-356160/37.
XX DR
```

DR WPI; 2001-356160/37.  
XX Polymorphic nucleic acid sequences, useful in genetic testing and  
PT therapy -  
XX  
XX  
PS Claim 29; Page 2641; 2653pp; English.  
XX  
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AAM53114 to AAM53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patients own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
XX  
SQ Sequence 1 AA;  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 X 1  
Db 1 L 1  
RESULT 15  
AAM53328  
ID AAM53328 standard; Peptide; 1 AA.  
XX  
XX  
AC AAM533328;  
XX  
DT 09-NOV-2001 (first entry)  
XX  
DE Human nonconservative amino acid changing SNP related peptide SEQ.7023.  
XX  
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
KW protein therapy; vaccine; probe; diagnostic assay; detection;  
KW quantitation; restorative therapy; polymorphic.  
XX  
OS Homo sapiens.  
XX  
PN WO200140521-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 30-NOV-2000; 2000WO-US32758.  
XX  
PR 30-NOV-1999; 99US-0168138.  
PR 29-NOV-2000; 2000US-0726173.  
XX  
XX (CURA-) CURAGEN CORP..  
PA  
PI Shimkets RA, Leach M;  
XX  
XX WPI; 2001-356160/37.  
DR  
XX Polymorphic nucleic acid sequences, useful in genetic testing and  
PT therapy -  
XX  
XX Claim 29; Page 2652; 2653pp; English.

XX  
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AAM53114 to AAM53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patients own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
XX  
SQ Sequence 1 AA;  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 X 1  
Db 1 V 1  
Search completed: December 8, 2003, 09:16:39  
Job time : 5.45455 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 4.20202 Seconds  
(without alignments)  
245.646 Million cell updates/sec

Title: US-09-498-556C-85  
Perfect score: 4  
Sequence: 1 XXXX 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rhodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	4	11 Q08433	Q08433 rattus sp.
2	0	0.0	5	2 P83073	P83073 bacillus ce
3	0	0.0	5	10 Q99007	Q99007 hordeum vul
4	0	0.0	5	13 P83308	P83308 gallus gall
5	0	0.0	6	10 P82181	P82181 spinacia ol
6	0	0.0	6	10 P82541	P82541 spinacia ol
7	0	0.0	6	10 P82182	P82182 spinacia ol
8	0	0.0	7	2 Q8KMS3	Q8KMS3 klebsiella
9	0	0.0	7	2 Q47505	Q47505 escherichia
10	0	0.0	7	2 P70804	P70804 azotobacter
11	0	0.0	7	2 O50556	O50556 actinobacil
12	0	0.0	7	2 Q47477	Q47477 escherichia
13	0	0.0	7	2 Q47029	Q47029 enterobacte
14	0	0.0	7	2 P72081	P72081 nocardia la
15	0	0.0	7	2 Q54248	Q54248 streptomyce
16	0	0.0	7	2 Q8KMS9	Q8KMS9 enterobacte

17	0	0.0	7	2	Q34028	Q34028 sphingomona
18	0	0.0	7	2	O07354	O07354 synechococc
19	0	0.0	7	2	Q8GL12	Q8GL12 borrelia bu
20	0	0.0	7	2	Q8GL04	Q8GL04 borrelia bu
21	0	0.0	7	2	Q8GL00	Q8GL00 borrelia bu
22	0	0.0	7	4	Q15903	Q15903 homo sapien
23	0	0.0	7	4	Q8NH7	Q8NH7 homo sapien
24	0	0.0	7	4	Q15897	Q15897 homo sapien
25	0	0.0	7	6	Q28742	Q28742 oryctolagus
26	0	0.0	7	8	P92214	P92214 amblyopyrum
27	0	0.0	7	8	P92393	P92393 hordeum vul
28	0	0.0	7	8	P92403	P92403 lophopyrum
29	0	0.0	7	8	P92427	P92427 peridictyon
30	0	0.0	7	8	O99182	O99182 gnatholebia
31	0	0.0	7	8	P92430	P92430 aegilops ta
32	0	0.0	7	8	P92221	P92221 bromus iner
33	0	0.0	7	8	O98866	O98866 spinacia ol
34	0	0.0	7	8	P92425	P92425 pseudoroegn
35	0	0.0	7	8	P92381	P92381 hordeum bra
36	0	0.0	7	8	P92387	P92387 henrardia p
37	0	0.0	7	8	P92210	P92210 agropyron c
38	0	0.0	7	8	P92440	P92440 thinopyrum
39	0	0.0	7	8	P92218	P92218 australopyr
40	0	0.0	7	8	P92390	P92390 heteranthel
41	0	0.0	7	8	P92372	P92372 haynaldia v
42	0	0.0	7	8	P92442	P92442 taeniatheru
43	0	0.0	7	8	P92226	P92226 crithopsis
44	0	0.0	7	8	Q8MFY6	Q8MFY6 taraxacum (
45	0	0.0	7	8	Q95945	Q95945 saccharomyc

## ALIGNMENTS

### RESULT 1

Q08433 PRELIMINARY; PRT; 4 AA.  
AC Q08433;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Bilirubin UDP-glucuronosyltransferase (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Gunn;  
RX MEDLINE=91282758; PubMed=1840486;  
RA Sato H., Aono S., Kashiwamata S., Koikai O.;  
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the  
hyperbilirubinemic Gunn rat."  
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).  
DR EMBL; S38636; AAB19259.1; -;  
KW Transferase.  
FT NON\_TER 1  
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 0.0%; Score 0; DB 11; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 N 1

### RESULT 2

P83073 PRELIMINARY; PRT; 5 AA.  
ID P83073  
AC P83073;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)

```

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE 88 kDa protein (Fragment).
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE.
RC STRAIN=NCIMB 11796;
RA Browne N., Dowds B.C.A.;
RL Submitted (JUL-2001) to the SWISS-PROT data bank.
FT NON TER 5
SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 5;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 4 T 4

RESULT 3
Q99007 PRELIMINARY; PRT; 5 AA.
AC Q99007;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha amylase (Fragment).
DE AMY1 GENE.
GN Hordeum vulgare (Barley).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91329704; PubMed=1831055;
RA Jacobsen J.V., Close T.J.;
RT "Control of transient expression of chimaeric genes by gibberellic
RT acid and abscisic acid in protoplasts prepared from mature barley
RT aleurone layers.";
RL Plant Mol. Biol. 16:713-721(1991).
DR EMBL; X54643; CAA38455.1; -.
FT NON TER 5
SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 5;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 A 2

RESULT 4
P83308 PRELIMINARY; PRT; 5 AA.
AC P83308;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE FMRamide-like neuropeptide (LPLRF-amide).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE, AND SYNTHESIS.

RC TISSUE=Brain;
RX PubMed=6137771;
RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
RT "A novel active pentapeptide from chicken brain identified by
RT antibodies to FMRamide.";
RL Nature 305:328-330(1983).
CC -|- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
CC -|- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide.
SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;

Query Match 0.0%; Score 0; DB 13; Length 5;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 L 1

RESULT 5
P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. ALMARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -|- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -|- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -|- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 A 1

RESULT 6
P82541 PRELIMINARY; PRT; 6 AA.
AC P82541;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS Spinacia oleracea (Spinach).

```

OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
 RX MEDLINE=20435797; PubMed=10874039;  
 RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 RL the small subunit of an organelle ribosome (chloroplast).";  
 RJ J. Biol. Chem. 275:28466-28482(2000).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -1- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.  
 CC -1- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.  
 CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA  
 CC FORM IS THE MINOR BASIC FORM.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro: IPR002222; Ribosomal S19.  
 DR Pfam: PF00203; Ribosomal S19; PARTIAL.  
 DR PRINTS; PR00975; RIBOSOMALS19; PARTIAL.  
 DR PROSITE; PS00323; RIBOSOMAL S19; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON TER 6 6  
 SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;  
 Query Match 0.0%; Score 0; DB 10; Length 6;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 1 T 1  
 RESULT 7  
 ID P82182 PRELIMINARY; PRT; 6 AA.  
 AC P82182;  
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L10 gamma (fragment).  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 RL the 50 S subunit of an organelle ribosome (chloroplast).";  
 RJ J. Biol. Chem. 275:28466-28482(2000).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro: IPR001790; Ribosomal L10.  
 DR Pfam: PF00466; Ribosomal L10; PARTIAL.  
 DR PROSITE; PS01109; RIBOSOMAL L10; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON TER 6 6  
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 1 A 1  
 RESULT 8  
 ID Q8KMS3 PRELIMINARY; PRT; 7 AA.  
 AC Q8KMS3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Putative Merr2 protein.  
 GN MERR2.  
 OS Klebsiella sp. LS13-39.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Klebsiella.  
 OX NCBI\_TaxID=143776;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LS13-39;  
 RX MEDLINE=21604134; PubMed=11763242;  
 RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,  
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,  
 RA Yurieva O.V., Nikiforov V.G.;  
 RT "Mercury resistance transposons of Gram-negative environmental  
 RT bacteria and their classification.";  
 RL Res. Microbiol. 152:811-822(2001).  
 DR EMBL; AJ302776; CAC82975.1; -;  
 SQ SEQUENCE 7 AA; 608 MW; 6DC1B5BDD87DD6F0 CRC64;  
 Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 2 A 2  
 RESULT 9  
 ID Q47505 PRELIMINARY; PRT; 7 AA.  
 AC Q47505;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE MccA protein.  
 GN MCCA.  
 OS Escherichia coli.  
 OG Plasmid pMccC7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96099297; PubMed=85252520;  
 RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;  
 RT "Structure and organization of plasmid genes required to produce the  
 RT translation inhibitor microcin C7.";  
 RL J. Bacteriol. 177:7131-7140(1995).  
 DR EMBL; X57583; CAA40808.1; -;  
 KW Plasmid.  
 SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;  
 Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 3 T 3

## RESULT 10

P70804 PRELIMINARY; PRT; 7 AA.  
AC P70804 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE AlqT protein (Fragment).  
GN ALGT.  
OS Azotobacter vinelandii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Azotobacter.  
OX NCBI\_TaxID=354;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E;  
RX MEDLINE=96427318; PubMed=8830682;  
RA Rehm B.H.A., Ertesvag H., Valla S.;  
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is  
RT part of an alg gene cluster physically organized in a manner similar  
RT to that in Pseudomonas aeruginosa.";  
RL J. Bacteriol. 178:5884-5889(1996).  
DR EMBL; X87973; CAA61230.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 2 T 2

## RESULT 11

O50556 PRELIMINARY; PRT; 7 AA.  
AC O50556;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GLYA (Fragment).  
GN GLYA.  
OS Actinobacillus actinomycetemcomitans (Haemophilus  
OS actinomycetemcomitans).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Actinobacillus.  
OX NCBI\_TaxID=714;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33384;  
RX MEDLINE=96355846; PubMed=8751884;  
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,  
RA Kraig E.;  
RT "cis Elements and trans factors are both important in strain-specific  
RT regulation of the leukotoxin gene in Actinobacillus  
RT actinomycetemcomitans.";  
RL Infect. Immun. 64:3451-3460(1996).  
DR EMBL; U51862; AAB88721.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 7 A 7

## RESULT 12

Q47477 PRELIMINARY; PRT; 7 AA.  
AC Q47477 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE Tpi protein (Fragment).  
GN TPI.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE OF 7-7 FROM N.A.  
RX MEDLINE=85203917; PubMed=3158524;  
RA Hellinga H.W., Evans P.R.;  
RT "Nucleotide sequence and high-level expression of the major  
RT Escherichia coli phosphofructokinase.";  
RL Eur. J. Biochem. 149:363-373(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Evans P.;  
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X02519; CAA26359.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 773 MW; 7416D33DDDBD1DB0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 A 1

## RESULT 13

Q47029 PRELIMINARY; PRT; 7 AA.  
AC Q47029 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Aad A1 protein (Fragment).  
GN Aad A1.  
OS Enterobacter cloacae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Enterobacter.  
OX NCBI\_TaxID=550;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94079349; PubMed=8257126;  
RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;  
RT "Analysis of the sac(3)-Via gene encoding a novel 3-N-  
RT acetyltransferase.";  
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).  
DR EMBL; M88012; AAA6193.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 2 T 2



## RESULT 14

P72081  
 ID P72081 PRELIMINARY; PRT; 7 AA.  
 AC P72081;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE 3'-methylcephem hydroxylase (Fragment).  
 GN CEFF.  
 OS Nocardia lactamdurans.  
 OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;  
 OC Pseudonocardiales; Pseudonocardaceae; Amycolatopsis.  
 OX NCBI\_TaxID=1913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96009872; PubMed=7557411;  
 RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,  
 RA Liras P.;  
 RT "Characterization of the cmh genes of Nocardia lactamdurans and  
 RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem  
 RT O-carbamoyltransferase for cephamycin biosynthesis.";  
 RL Gene 162:21-27(1995).  
 DR EMBL; Z21682; CAA79797.1; --  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 T 1

## RESULT 15

Q54248  
 ID Q54248 PRELIMINARY; PRT; 7 AA.  
 AC Q54248;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE RpLO protein (Fragment).  
 GN RpLO.  
 OS Streptomyces griseus.  
 OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=N2-3-11;  
 RX MEDLINE=20011291; PubMed=10542330;  
 RA Poehling S., Piepersberg W., Wehmeier U.F.;  
 RT "Analysis and regulation of the sec Y gene from Streptomyces griseus  
 RT N2-3-11 and interaction of the SecY protein with the SecA protein.";  
 RL Biochim. Biophys. Acta 1447:298-302(1999).  
 DR EMBL; X95915; CAA65160.1; --  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 2 T 2

Search completed: December 8, 2003, 09:18:40  
 Job time : 5.20202 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.05051 Seconds  
(without alignments)  
179.063 Million cell updates/sec

Title: US-09-498-556C-85  
Perfect score: 4  
Sequence: 1 XXXX 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	ID	Description
1	0	0.0	3 1 GRWM HUMAN	P01157 homo sapien
2	0	0.0	3 1 LUXE_VIBFI	P24272 vibrio fisc
3	0	0.0	3 1 THYL_PIG	P01151 sus scrofa
4	0	0.0	4 1 ACHI_ACHFU	P35904 achatina fu
5	0	0.0	4 1 DCML_PSECH	P19916 pseudomonas
6	0	0.0	4 1 DCMS_PSECH	P19918 pseudomonas
7	0	0.0	4 1 EOST_HUMAN	P02731 homo sapien
8	0	0.0	4 1 FAR3_HIRME	P42562 hirudo medi
9	0	0.0	4 1 FAR4_HIRME	P42563 hirudo medi
10	0	0.0	4 1 FLRF_HIRME	P58705 anthopleura
11	0	0.0	4 1 FLRF_HIRME	P42561 hirudo medi
12	0	0.0	4 1 FLRN_ATEL	P58707 anthopleura
13	0	0.0	4 1 FMRF_WACNI	P01162 macrocallis
14	0	0.0	4 1 FYRI_ATEL	P58706 anthopleura
15	0	0.0	4 1 OCPI_OCTMI	P58648 octopus min
16	0	0.0	4 1 OCPI_OCTMI	P58649 octopus min
17	0	0.0	4 1 RM01_YEAST	P36515 saccharomyc
18	0	0.0	4 1 TUFT_HUMAN	P01858 homo sapien
19	0	0.0	5 1 ALI4_CARMA	P81817 carcinus ma
20	0	0.0	5 1 B10A_CITFR	P13071 citrobacter
21	0	0.0	5 1 B10B_CITFR	P12997 citrobacter
22	0	0.0	5 1 BPP7_BOTIN	P30425 bothrops in
23	0	0.0	5 1 EI03_LITRU	P82099 litoria rub
24	0	0.0	5 1 EI04_LITRU	P82100 litoria rub
25	0	0.0	5 1 FARP_ARTTR	P41853 artiopesthi
26	0	0.0	5 1 PAP2_FARMA	P81864 pardachirus
27	0	0.0	5 1 PRCT_PERAM	P01373 periplaneta
28	0	0.0	5 1 PSK_DAUCA	P58261 daucus caro
29	0	0.0	5 1 RE11_LITRU	P82070 litoria rub
30	0	0.0	5 1 RE21_LITRU	P82071 litoria rub
31	0	0.0	5 1 RE31_LITRU	P82072 litoria rub
32	0	0.0	5 1 RE32_LITRU	P82073 litoria rub
33	0	0.0	5 1 SUGA_ACHDO	P19991 acheta dome

## ALIGNMENTS

## RESULT 1

GRWM HUMAN STANDARD; PRT; 3 AA.  
 AC P01157;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DE Growth-modulating peptide.  
 OS Homo sapiens (Human);  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77162369; PubMed=858356;  
 RA Schlesinger D.H., Pickart L., Thaler M.M.;  
 RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";  
 RL Experientia 33:324-325(1977).  
 CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE  
 CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.  
 DR GO: GO:0001558; P:regulation of cell growth; NAS.  
 SQ SEQUENCE 3 AA; 340 MW; 6331E91000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 G 1

## RESULT 2

LUXE\_VIBFI STANDARD; PRT; 3 AA.  
 ID LUXE\_VIBFI  
 AC P24272;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Long-chain-fatty-acid-luciferin-component ligase (EC 6.2.1.19) (Acyl-protein synthetase) (Fragment).  
 DE LUXE.  
 GN Vibrio fischeri.  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=668;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91072226; PubMed=2254256;  
 RA Swartzman E., Kapoor S., Graham A.F., Meighen B.A.;  
 RT "A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon.";  
 RL J. Bacteriol. 172:6797-6802(1990).  
 CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.  
 CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS



RA Iwashita T., Nomoto K.;  
 RT "Crystal structure and molecular conformation of achainin-I  
 (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a  
 D-amino acid residue.";  
 RL Int. J. Pept. Protein Res. 39:258-264(1992).  
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY  
 AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY  
 NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE  
 HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.  
 CC PIR; A32480; A32480.  
 DR Hormone; D-amino acid.  
 KW MOD\_RES 2 2 D-PHENYLALANINE.  
 FT MOD\_RES 4 AA; 408 MW; 6AAD9C8100000000 CRC64;  
 SQ SEQUENCE 4 AA; 408 MW; 6AAD9C8100000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 0; Indels 1; Gaps 0;  
 QY 1 X 1  
 Db 3 A 3  
 RESULT 5  
 DCML\_PSECH STANDARD; PRT; 4 AA.  
 ID DCML\_PSECH  
 AC P19916;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO  
 dehydrogenase subunit L) (CO-DH L) (Fragment).  
 GN CUTL.  
 OS Pseudomonas carboxydohydrogena.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae.  
 OX NCBI\_TaxID=290;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90055678; PubMed=2818128;  
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
 RT "Homology and distribution of CO dehydrogenase structural genes in  
 carboxydohydrogenic bacteria.";  
 RL Arch. Microbiol. 152:335-341(1989).  
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon  
 dioxide.  
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced  
 acceptor.  
 CC -1- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.  
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
 SMALL.  
 CC PIR; PLO146; PLO146.  
 KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.  
 FT NON\_TER 4  
 SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F00000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 2 A 2  
 RESULT 7  
 EOSI\_HUMAN STANDARD; PRT; 4 AA.  
 ID EOSI\_HUMAN  
 AC P02731;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DE Eosinophilotoxic peptides.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76078412; PubMed=1060093;  
 RA Goetzl E.J., Austen K.F.;  
 RT "Purification and synthesis of eosinophilotoxic tetrapeptides of  
 human lung tissue: Identification as eosinophil chemotactic factor of  
 anaphylaxis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).  
 CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG  
 (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS  
 (ANAPHYLAXIS). THEIR ACTIVITIES, PREPOTENTIALLY AFFECTING  
 EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE  
 OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.  
 DR GO; GO:0030105; P:anaphylaxis; IDA.  
 DR GO; GO:0006935; P:chemotaxis; IDA.  
 FT VARIANT 1 1 V -> A (IN OTHER PEPTIDE).  
 FT /FTID=VAR 005201.  
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A00000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 3 S 3  
 RESULT 6  
 DCMS\_PSECH STANDARD; PRT; 4 AA.  
 ID DCMS\_PSECH  
 AC P19918;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO

DE dehydrogenase subunit S) (CO-DH S) (Fragment).  
 GN CUTS.  
 OS Pseudomonas carboxydohydrogena.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae.  
 OX NCBI\_TaxID=290;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90055678; PubMed=2818128;  
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
 RT "Homology and distribution of CO dehydrogenase structural genes in  
 carboxydohydrogenic bacteria.";  
 RL Arch. Microbiol. 152:335-341(1989).  
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon  
 dioxide.  
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced  
 acceptor.  
 CC -1- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.  
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
 SMALL.  
 CC PIR; PLO146; PLO146.  
 KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.  
 FT NON\_TER 4  
 SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F00000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 2 A 2  
 RESULT 7  
 EOSI\_HUMAN STANDARD; PRT; 4 AA.  
 ID EOSI\_HUMAN  
 AC P02731;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DE Eosinophilotoxic peptides.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76078412; PubMed=1060093;  
 RA Goetzl E.J., Austen K.F.;  
 RT "Purification and synthesis of eosinophilotoxic tetrapeptides of  
 human lung tissue: Identification as eosinophil chemotactic factor of  
 anaphylaxis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).  
 CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG  
 (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS  
 (ANAPHYLAXIS). THEIR ACTIVITIES, PREPOTENTIALLY AFFECTING  
 EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE  
 OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.  
 DR GO; GO:0030105; P:anaphylaxis; IDA.  
 DR GO; GO:0006935; P:chemotaxis; IDA.  
 FT VARIANT 1 1 V -> A (IN OTHER PEPTIDE).  
 FT /FTID=VAR 005201.  
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A00000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 3 S 3

## RESULT 8

FAR3\_HIRME STANDARD; PRT; 4 AA.  
AC P42562;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FMRamide-like neuropeptide YNRF-amide.  
OS Hirudo medicinalis (Medicinal leech).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92195954; PubMed=1686933;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of Rfamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD RES 4  
SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Y 1

## RESULT 9

FAR4\_HIRME STANDARD; PRT; 4 AA.  
AC P42563;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FMRamide-like neuropeptide YNRF-amide.  
OS Hirudo medicinalis (Medicinal leech).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92195954; PubMed=1686933;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of Rfamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD RES 4  
SQ SEQUENCE 4 AA; 616 MW; 69D4068B300000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Y 1

## RESULT 10

FFKA\_ATEL STANDARD; PRT; 4 AA.  
AC P58705;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Antho-Kaamide.  
OS Anthopleura elegantissima (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
OC Nynanthaeae; Actiniidae; Anthopleura.  
OX NCBI\_TaxID=6110;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92028852; PubMed=1681803;  
RA Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;  
RT "Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-Kaamide), a  
novel neuropeptide from sea anemones.";  
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).  
RN [2]  
RP FUNCTION.

RX MEDLINE=93391436; PubMed=8397415;  
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;  
RT "The expansion behaviour of sea anemones may be coordinated by two  
inhibitory neuropeptides, Antho-Kaamide and Antho-Riamide.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).  
CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle  
groups. May be involved in the expansion phase of feeding  
behaviour in sea anemones.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Neuron-specific.  
DR PIR; JQ1273; JQ1273.  
KW Neuropeptide; Amidation.  
FT MOD RES 1 1 L-3-PHENYLLACTYL.  
FT MOD RES 4 4 AMIDATION.  
SQ SEQUENCE 4 AA; 512 MW; 6DD339C9A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 4 A 4

## RESULT 11

FLRF\_HIRME STANDARD; PRT; 4 AA.  
AC P42561;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FLRFamide.  
OS Hirudo medicinalis (Medicinal leech), and  
OS Helisoma trivolvis (Snail).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421, 27815;  
RN [1]  
RP SEQUENCE.  
RX SPECIES=H.medicalinalis;  
RX MEDLINE=92195954; PubMed=1686933;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of Rfamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
RN [2]  
RP SEQUENCE.

RX SPECIES=H.trivolvis; TISSUE=Kidney;  
RX MEDLINE=94286417; PubMed=7912428;  
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
trivolvis.";  
RL Peptides 15:31-36(1994).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD RES 4 4 AMIDATION.  
SQ SEQUENCE 4 AA; 582 MW; 69D40729A00000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 F 1

## RESULT 12

FLRN ANTEL STANDARD; PRT; 4 AA.  
 AC P58707;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Antho-RNamide.  
 OS Anthopleura elegantissima (Sea anemone).  
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
 OC Nynanthaeae; Actiniidae; Anthopleura.  
 OX NCBI\_TaxID=6110;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RX MEDLINE=90319122; PubMed=1973541;  
 RA Grimmlikhuizen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,  
 RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;  
 RT "Isolation of L-3-phenylacetyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea  
 RT anemone neuropeptide containing an unusual amino-terminal blocking  
 RT group.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414 (1990).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Neuron-specific.  
 CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.  
 DR PIR; A35779; A35779.  
 KW Neuropeptide; Amidation.  
 FT MOD RES 1 1 L-3-PHENYLACTYL.  
 FT MOD RES 4 4 AMIDATION.  
 SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 F 1

## RESULT 13

FMRF MACNI STANDARD; PRT; 4 AA.  
 AC P01162;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE FMRamide (peak C) (Cardioexcitatory neuropeptide).  
 OS Macrocallista nimbosa (Sun-ray clam),  
 OS Nereis virens (Sandworm),  
 OS Hirudo medicinalis (Medicinal leech), and  
 OS Helisoma trivolvis (Snail).  
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;  
 OC Veneroidea; Veneridae; Macrocallista.  
 OX NCBI\_TaxID=6594, 6353, 6421, 27815,  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;  
 RX MEDLINE=77215956; PubMed=877582;  
 RA Price D.A., Greenberg M.J.;  
 RT "Structure of a molluscan cardioexcitatory neuropeptide.";  
 RL Science 197:670-671 (1977).  
 RN [2]  
 RP SEQUENCE, AND CHARACTERIZATION.

RC SPECIES=M.nimbosa; TISSUE=Ganglion;  
 RX MEDLINE=78012038; PubMed=909875;  
 RA Price D.A., Greenberg M.J.;  
 RT "Purification and characterization of a cardioexcitatory neuropeptide  
 RT from the central ganglia of a bivalve mollusc.";  
 RL Prep. Biochem. 7:261-281 (1977).  
 RN [3]  
 RP SEQUENCE.  
 RC SPECIES=N.virens;  
 RX MEDLINE=90259866; PubMed=2342992;  
 RA Krajinak K.G., Price D.A.;  
 RT "Authentic FMRamide is present in the polychaete Nereis virens.";  
 RL Peptides 11:75-77 (1990).  
 RN [4]  
 RP SEQUENCE.  
 RC SPECIES=H.medicinalis;  
 RX MEDLINE=92195954; PubMed=1686933;  
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
 RT "Identification of RFamide neuropeptides in the medicinal leech.";  
 RL Peptides 12:897-908 (1991).  
 RN [5]  
 RP SEQUENCE.  
 RC SPECIES=H.trivolvis; TISSUE=Kidney;  
 RX MEDLINE=94286417; PubMed=7912428;  
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
 RT "FMRamide-related peptides from the kidney of the snail, Helisoma  
 RT trivolvis";  
 RL Peptides 15:31-36 (1994).  
 CC -!- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL  
 CC ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULARIZATION OF  
 CC CARDIAC CONTRACTION.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR; A01426; ECNK.  
 DR PIR; A60418; A60418.  
 KW Neuropeptide; Amidation.  
 FT MOD RES 4 4 AMIDATION.  
 SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 F 1

## RESULT 14

FYRI ANTEL STANDARD; PRT; 4 AA.  
 AC P58706;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Antho-RNamide I [Contains: Antho-RNamide II].  
 OS Anthopleura elegantissima (Sea anemone).  
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
 OC Nynanthaeae; Actiniidae; Anthopleura.  
 OX NCBI\_TaxID=6110;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92270459; PubMed=1821096;  
 RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,  
 RA Grimmlikhuizen C.J.P.;  
 RT "Isolation of two novel  
 RT biologically active L-3-phenylacetyl-Tyr-Arg-Ile-NH2 and its  
 RT des-phenylacetyl fragment Tyr-Arg-Ile-NH2.";  
 RL Peptides 12:1165-1173 (1991).  
 RN [2]  
 RP FUNCTION.  
 RX MEDLINE=93391436; PubMed=8397415;  
 RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmlikhuizen C.J.P.;

RT "The expansion behaviour of sea anemones may be coordinated by two  
RT inhibitory neuropeptides, Antho-Kamide and Antho-Riamide.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).  
CC -I- FUNCTION: Inhibits spontaneous contractions in several muscle  
CC groups. May be involved in the expansion phase of feeding  
CC behaviour in sea anemones.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- TISSUE SPECIFICITY: Neuron-specific.  
KW Neuropeptide; Amidation.  
FT CHAIN 1 4 ANTHO-RIAMIDE I.  
FT CHAIN 2 4 ANTHO-RIAMIDE II.  
FT MOD RES 1 1 L-3-PHENYLLACTYL.  
FT MOD RES 4 4 AMIDATION.  
SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 F 1

## RESULT 15

OCPL OCTMI STANDARD; PRT; 4 AA.  
AC P58648;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cardioactive peptides Ocp-1/Ocp-2.  
OS Octopus minor (Octopus).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.  
OX NCBI\_TaxID=89766;  
RN [1]  
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=20336815; PubMed=10876044;  
RA Iwakoshi E., Hisada M., Minakata H.;  
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
RT Octopus minor.";  
RL Peptides 21:623-630(2000).  
CC -I- FUNCTION: Cardioactive; has both positive chronotropic and  
CC inotropic effects on the heart. Ocp-2 is a 1000 time less  
CC active than Ocp-1.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- PTM: Ocp-2 has L-Phe instead of D-Phe.  
CC -I- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
KW Hormone; D-amino acid.  
FT MOD RES 2 2 D-PHENYLALANINE.  
SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 G 1

Search completed: December 8, 2003, 09:14:14  
Job time : 2.05051 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.65657 Seconds  
(without alignments)  
232.212 Million cell updates/sec

Title: US-09-498-556C-85

Perfect score: 4  
Sequence: 1 XXXX 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	3 RHTDTC	thyroliberin - Bom
2	0	0.0	3	3 RHPCT	thyroliberin - pig
3	0	0.0	3	3 RSHST	thyroliberin - she
4	0	0.0	3	3 A92971	thyroliberin - eas
5	0	0.0	3	3 GKHU	growth-modulating
6	0	0.0	3	3 A60898	bursin - chicken
7	0	0.0	3	3 A23751	spinal cord peptid
8	0	0.0	3	3 B23751	spinal cord peptid
9	0	0.0	3	3 A23802	thyrotropin-releas
10	0	0.0	3	3 A22565	R-phycoerythrin al
11	0	0.0	3	3 PQ0010	angiotensin-conver
12	0	0.0	3	3 S13894	histidinol dehydro
13	0	0.0	3	3 A43391	TRH-like tripeptid
14	0	0.0	3	3 E37196	bradykinin-potenti
15	0	0.0	3	3 F37196	bradykinin-potenti
16	0	0.0	3	3 I50412	gene p20K protein
17	0	0.0	3	3 PT0636	T-cell receptor be
18	0	0.0	3	3 PT0578	T-cell receptor be
19	0	0.0	3	3 PT0571	T-cell receptor be
20	0	0.0	3	3 PT0622	T-cell receptor be
21	0	0.0	3	3 I78890	tyrosine protein k
22	0	0.0	3	3 S68328	blood cell protein
23	0	0.0	3	3 T13892	cytochrome-c oxida
24	0	0.0	4	1 ECXAA	antho-Ramide neur
25	0	0.0	4	2 S18401	thyroglobulin - do
26	0	0.0	4	2 A02147	phagocytosis-stimu
27	0	0.0	4	2 A32039	tyrosine-melanocyt
28	0	0.0	4	2 ECNK	cardioexcitatory n
29	0	0.0	4	2 PL0140	carbon-monoxide de

30 0 0.0 4 2 PL0146 carbon-monoxide de  
31 0 0.0 4 2 A37832 phenol 2-monooxyge  
32 0 0.0 4 2 A48360 gamma subunit of p  
33 0 0.0 4 2 I40697 biotin A - Citroba  
34 0 0.0 4 2 A61300 22K superhelical D  
35 0 0.0 4 2 I57745 D-mannonate hydrol  
36 0 0.0 4 2 A41890 protein D - Escher  
37 0 0.0 4 2 S43014 hypothetical prote  
38 0 0.0 4 2 D41654 hypothetical prote  
39 0 0.0 4 2 B43848 cell surface adhes  
40 0 0.0 4 2 I40505 hypothetical prote  
41 0 0.0 4 2 I40870 phospholipase C (E  
42 0 0.0 4 2 I40804 endoglucanase F -  
43 0 0.0 4 2 T46627 starvation-induced  
44 0 0.0 4 2 S53508 ribosomal protein  
45 0 0.0 4 2 S17255

#### ALIGNMENTS

##### RESULT 1

RHTDTC

thyroliberin - Bombina orientalis

C:Species: Bombina orientalis

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A90919; A01415

R:Yasuhara, T.; Nakajima, T.

Chem. Pharm. Bull. 23, 3301-3303, 1975

A:Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.

A:Reference number: A90919; MUID:76138399; PMID:815011

A:Accession: A90919

A:Molecule type: protein

A:Residues: 1-3 <YAS>

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Q 1

##### RESULT 2

RHPGT

thyroliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A01415

R:Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.

Biochemistry 9, 1103-1106, 1970

A:Title: Structure of porcine thyrotropin releasing hormone.

A:Reference number: A90560; MUID:70136150; PMID:4984938

A:Accession: A01415

A:Molecule type: protein

A:Residues: 1-3 <NAI>

R:Boiler, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.

Biochem. Biophys. Res. Commun. 37, 705-710, 1969

A:Title: The identity of chemical and hormonal properties of the thyrotropin releasing

A:Reference number: A90167; MUID:70039904; PMID:4982117

A:Contents: annotation

A>Note: biological activities and Rf values (in 17 chromatographic systems) of the synt

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 3  
RHSHY  
thyroliberin - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A93750; A01415  
R:Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.  
Org. Mass Spectrom. 5, 221-228, 1971  
A:Title: The elucidation of the primary structure of the hypothalamic thyroïd stimulating hormone  
A:Reference number: A93750  
A:Accession: A93750  
A:Molecule type: protein  
A:Residues: 1-3 <DES>  
R:Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.  
Nature 226, 321-325, 1970  
A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.  
A:Reference number: A93161; MUID:70163386; PMID:4985794  
A:Contents: annotation  
A:Note: physicochemical characteristics and biological activities of the natural and synthetic thyroliberin precursor  
C:Superfamily: thyroliberin precursor  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 4  
A92971  
thyroliberin - eastern newt (tentative sequence)  
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A92971; A01415  
R:Grimm-Jorgensen, Y.; McKelvy, J.F.  
J. Neurochem. 23, 471-478, 1974  
A:Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain  
A:Reference number: A92971; MUID:75035605; PMID:4214528  
A:Accession: A92971  
A:Molecule type: protein  
A:Residues: 1-3 <GRI>  
A:Note: a peptide with the chromatographic and electrophoretic characteristics of thyrotropin, or glutamic acid  
C:Superfamily: thyroliberin precursor  
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 5  
GKHU  
growth-modulating peptide - human

C:Species: Homo sapiens (man)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A01421  
R:Schlesinger, D.H.; Pickart, L.; Thaler, M.M.  
Experientia 33, 324-325, 1977  
A:Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.  
A:Reference number: A01421, MUID:77162369; PMID:858356  
A:Accession: A01421  
A:Molecule type: protein  
A:Residues: 1-3 <SCH>  
A:Note: this serum tripeptide is found to stimulate growth of some cell types and to inhibit growth of others  
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 G 1

RESULT 6  
A60898  
bursin - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A60898  
R:Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.  
Science 231, 997-999, 1986  
A:Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone of the bursa of Fabricius  
A:Reference number: A60898; MUID:86122916; PMID:3484838  
A:Accession: A60898  
A:Molecule type: protein  
A:Residues: 1-3 <AUD>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; hormone  
F:3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 K 1

RESULT 7  
A23751  
spinal cord peptide SCP-4 - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A23751  
R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
Arch. Biochem. Biophys. 240, 178-183, 1985  
A:Reference number: A23751; MUID:85250425; PMID:4015098  
A:Accession: A23751  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <HSI>  
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 8

B23751  
 spinal cord peptide SCP-5 - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C:Accession: B23751  
 R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
 Arch. Biochem. Biophys. 240, 178-183, 1985  
 A:Reference number: A23751; MUID:85250425; PMID:4015098  
 A:Accession: B23751  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-3 <HSI>  
 C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 M 1

RESULT 9  
 A33802  
 thyrotropin-releasing hormone-like peptide - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C:Accession: A33802  
 R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.  
 J. Biol. Chem. 264, 7788-7791, 1989  
 A:Title: A novel peptide, pyroglutamylglutamyloproline amide, in the rabbit prostate comp  
 A:Reference number: A33802; MUID:89255196; PMID:2498305  
 A:Accession: A33802  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-3 <COC>  
 C:Superfamily: unassigned animal peptides  
 C:Keywords: amidated carboxyl end; pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 Q 1

RESULT 10  
 A22565  
 R-phycoerythrin alpha-1 chain - red alga (Gastrocloonium coulteri) (fragment)  
 C:Species: Gastrocloonium coulteri  
 C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C:Accession: A22565  
 R:Klotz, A.V.; Glazer, A.N.  
 J. Biol. Chem. 260, 4856-4863, 1985  
 A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
 A:Reference number: A22565; MUID:85182601; PMID:3886644  
 A:Accession: A22565  
 A:Molecule type: protein  
 A:Residues: 1-3 <KLO>

Query Match 0.0%; Score 0; DB 3; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 2 Y 2

RESULT 11  
 PQ0010  
 angiotensin-converting enzyme inhibitor (FLP-3) - common fig  
 N:Alternate names: ficus latex peptide 3  
 C:Species: Ficus carica (common fig)  
 C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C:Accession: PQ0010  
 R:Maruyama, S.; Miyoshi, S.; Tanaka, H.  
 Agric. Biol. Chem. 53, 2763-2767, 1989  
 A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.  
 A:Reference number: PQ0008  
 A:Accession: PQ0010  
 A:Molecule type: protein  
 A:Residues: 1-3 <MAR>  
 A:Experimental source: latex  
 C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 0.0%; Score 0; DB 3; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 L 1

RESULT 12  
 S13894  
 histidinol dehydrogenase (EC 1.1.1.23) - wild cabbage (fragment)  
 C:Species: Brassica oleracea (wild cabbage)  
 C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C:Accession: S13894  
 R:Nagai, A.; Scheidegger, A.  
 Arch. Biochem. Biophys. 284, 127-132, 1991  
 A:Title: Purification and characterization of histidinol dehydrogenase from cabbage.  
 A:Reference number: S13894; MUID:91112783; PMID:1989490  
 A:Accession: S13894  
 A:Molecule type: protein  
 A:Residues: 1-3 <NAG>  
 A:Experimental source: var. capitata  
 C:Keywords: dimer; NAD; oxidoreductase

Query Match 0.0%; Score 0; DB 3; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 X 1

RESULT 13  
 A43391  
 TRH-like tripeptide - alfalfa  
 C:Species: Medicago sativa (alfalfa)  
 C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C:Accession: A43391  
 R:Lackey, D.B.  
 J. Biol. Chem. 267, 17508-17511, 1992  
 A:Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroglu-  
 A:Reference number: A43391; MUID:92388092; PMID:1517203  
 A:Accession: A43391  
 A:Molecule type: protein  
 A:Residues: 1-3 <LAC>  
 C:Keywords: amidated carboxyl end; pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

## RESULT 14

E37196  
bradykinin-potentiating peptide 5 - island jararaca  
C;Species: Bothrops insularis (island jararaca)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 30-Jun-2001  
C;Accession: E37196  
R;Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.  
J. Protein Chem. 9, 221-227, 1990  
A;Title: Primary structure and biological activity of bradykinin potentiating peptides  
A;Reference number: A37196; MUID:90351557; PMID:2386615  
A;Accession: E37196  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-3 <CIN>  
C;Keywords: pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

## RESULT 15

F37196  
bradykinin-potentiating peptide 6 - island jararaca  
C;Species: Bothrops insularis (island jararaca)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: F37196  
R;Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.  
J. Protein Chem. 9, 221-227, 1990  
A;Title: Primary structure and biological activity of bradykinin potentiating peptides  
A;Reference number: A37196; MUID:90351557; PMID:2386615  
A;Accession: F37196  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-3 <CIN>  
C;Keywords: pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

Search completed: December 8, 2003, 09:19:33  
Job time : 2.65657 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 1.85859 Seconds  
(without alignments)  
91.060 Million cell updates/sec

Title: US-09-498-556C-357  
Perfect score: 11  
Sequence: 1 LXPX 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCFUS.COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	3	1	US-08-470-837-7
2	9	81.8	3	4	US-08-868-452-7
3	9	81.8	4	1	US-08-288-405A-4
4	9	81.8	4	1	US-08-336-343A-36
5	9	81.8	4	3	US-09-142-078-19
6	9	81.8	4	3	US-09-357-141-19
7	9	81.8	4	4	US-09-533-889-19
8	9	81.8	4	4	US-09-142-080-19
9	9	81.8	4	4	US-09-486-283C-4
10	9	81.8	5	1	US-07-626-923A-12
11	9	81.8	5	1	US-08-357-264-5
12	9	81.8	5	1	US-08-456-840-38
13	9	81.8	5	1	US-08-672-514-5
14	9	81.8	5	1	US-07-789-184-173
15	9	81.8	5	1	US-08-406-192-12
16	9	81.8	5	1	US-08-460-343B-37
17	9	81.8	5	1	US-08-460-343B-38
18	9	81.8	5	1	US-08-475-263-173
19	9	81.8	5	1	US-08-398-028B-37
20	9	81.8	5	1	US-08-398-028B-38
21	9	81.8	5	1	US-08-266-407A-38
22	9	81.8	5	1	US-08-485-886-173
23	9	81.8	5	2	US-08-504-265B-37
24	9	81.8	5	2	US-08-504-265B-38
25	9	81.8	5	2	US-08-545-151-12
26	9	81.8	5	2	US-08-392-973A-7
27	9	81.8	5	2	US-08-392-973A-8

28	9	81.8	5	2	US-08-392-973A-9
29	9	81.8	5	2	US-08-477-362-173
30	9	81.8	5	2	US-08-477-134-173
31	9	81.8	5	2	US-08-892-544-38
32	9	81.8	5	3	US-08-473-489A-173
33	9	81.8	5	3	US-08-485-695-173
34	9	81.8	5	3	US-08-018-760-173
35	9	81.8	5	3	US-09-246-500B-17
36	9	81.8	5	4	US-09-367-791A-108
37	9	81.8	5	6	5169933-16
38	9	81.8	5	6	5464756-42
39	9	81.8	6	1	US-08-236-427-11
40	9	81.8	6	1	US-07-923-724-35
41	9	81.8	6	2	US-08-631-427A-3
42	9	81.8	6	2	US-08-806-203-2
43	9	81.8	6	2	US-08-806-203-3
44	9	81.8	6	2	US-08-806-203-4
45	9	81.8	6	2	US-08-806-203-5

## ALIGNMENTS

RESULT 1  
US-08-470-837-7  
; Sequence 7, Application US/08470837  
; Patent No. 5800811  
; GENERAL INFORMATION:  
; APPLICANT: Nimmi, Marcel E.  
; APPLICANT: Hall, Frederick L.  
; APPLICANT: Tuan, Tai-Lan  
; APPLICANT: Wu, Lingtao  
; APPLICANT: Cheung, David T.  
; TITLE OF INVENTION: Transforming Growth Factor B Fusion  
; TITLE OF INVENTION: and  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant & Gould  
; STREET: 11150 Santa Monica Boulevard, Suite 400  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90025-3395  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,837  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Janice A.  
; REGISTRATION NUMBER: 34,051  
; REFERENCE/DOCKET NUMBER: 30630-1US01  
; TELEPHONE: 310-445-1140  
; TELEFAX: 310-445-9031  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-470-837-7

Sequence 9, Appli  
Sequence 173, App  
Sequence 173, App  
Sequence 38, Appl  
Sequence 173, App  
Sequence 173, App  
Sequence 17, Appl  
Sequence 108, App  
Patent No. 5169933  
Patent No. 5464756  
Sequence 11, Appl  
Sequence 35, Appl  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 4, Appli  
Sequence 5, Appli

Query Match 81.8%; Score 9; DB 1; Length 3;  
Best Local Similarity 66.7%; Pred. NO. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

[illegible]

Sequence 19, Application US/09142078  
Patent No. 6172041  
GENERAL INFORMATION:  
APPLICANT: McCabe, R. Tyler  
APPLICANT: Zhou, Li-Ming  
APPLICANT: Layer, Richard T.  
TITLE OF INVENTION: Use of Conantokins  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701-E  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/142,078  
FILING DATE: 10-FEB-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO US97/12652  
FILING DATE: 21-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/762,377  
FILING DATE: 06-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/684,750  
FILING DATE: 22-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 2314-135.A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 4  
OTHER INFORMATION: /note= "Xaa is gamma-carboxyglutamic acid"  
US-09-142-078-19

Query Match 81.8%; Score 9; DB 3; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
Db 1 LAR 3

RESULT 6  
US-09-357-141-19  
Sequence 19, Application US/09357141  
Patent No. 6277825  
GENERAL INFORMATION:  
APPLICANT: Olivera, Baldomero M.  
APPLICANT: McIntosh, J. Michael  
APPLICANT: McCabe, R. Tyler  
APPLICANT: Layer, Richard T.  
APPLICANT: Zhou, Li-Ming

TITLE OF INVENTION: Use of Conantokins for Treating Pain  
FILE REFERENCE: 2314-171  
CURRENT APPLICATION NUMBER: US/09/357,141  
CURRENT FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: US 09/283,277  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: US 09/142,078  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: WO US97/12652  
PRIOR FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: US 08/762,377  
PRIOR FILING DATE: 1996-12-06  
PRIOR APPLICATION NUMBER: US 08/684,750  
PRIOR FILING DATE: 1996-07-22  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 19  
LENGTH: 4  
TYPE: PRT  
ORGANISM: Conus radiatus  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (4)  
OTHER INFORMATION: Xaa is gamma-carboxyglutamic acid.  
US-09-357-141-19

Query Match 81.8%; Score 9; DB 3; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
Db 1 LAR 3

RESULT 7  
US-09-533-889-19  
Sequence 19, Application US/09533889  
Patent No. 639574  
GENERAL INFORMATION:  
APPLICANT: McCabe, R. Tyler  
APPLICANT: Zhou, Li-Ming  
APPLICANT: Layer, Richard T.  
APPLICANT: Olivera, Baldomero M.  
APPLICANT: McIntosh, J. Michael  
TITLE OF INVENTION: Use of Conantokins  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701-E  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/533,889  
FILING DATE: 22 MAR-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/142,078  
FILING DATE: 10-FEB-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO US97/12652  
FILING DATE: 21-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/762,377  
FILING DATE: 06-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/684,750

```
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Innen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-168.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note= "Xaa is
; gamma-carboxyglutamic acid"
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-533-889-19

Query Match 81.8%; Score 9; DB 4; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LAR 3

RESULT 8
US-09-142-080-19
; Sequence 19, Application US/09142080
; Patent No. 6515103
; GENERAL INFORMATION:
; APPLICANT: Abogadie, Fe C.
; Cruz, Lourdes J.
; Olivera, Baldomero M.
; Walker, Craig
; Colledge, Clark
; Hillyard, David R.
; Jimenez, Elsie
; Laver, Richard T.
; Zhou, Li-Ming
; McCabe, R. Tyler
; TITLE OF INVENTION: Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figs, Earnet & Manbeck, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,080
; FILING DATE: 11-May-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12618
; FILING DATE: 21-JUL-1997
; APPLICATION NUMBER: US 08/684,742
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Innen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-134.A
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note= "Xaa is
; gamma-carboxyglutamic acid"
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-142-080-19

Query Match 81.8%; Score 9; DB 4; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LAR 3

RESULT 9
US-09-486-283C-4
; Sequence 4, Application US/09486283C
; Patent No. 6573243
; GENERAL INFORMATION:
; APPLICANT: ITAGAKI, Yasuhiro
; APPLICANT: KONNO, Katsuhiko
; APPLICANT: KAWAI, No. 6573243ufumi
; APPLICANT: TAKAYAMA, Hiroaki
; TITLE OF INVENTION: POMPILID WASP-DERIVED NEUROPEPTIDES
; FILE REFERENCES: 1830/48652
; CURRENT APPLICATION NUMBER: US/09/486,283C
; CURRENT FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: JP 9/241699
; PRIOR FILING DATE: 1997-08-25
; PRIOR APPLICATION NUMBER: PCT/JP98/03730
; PRIOR FILING DATE: 1998-08-24
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Pompilid sp.
; US-09-486-283C-4

Query Match 81.8%; Score 9; DB 4; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LSR 3

RESULT 10
US-07-626-923A-12
; Sequence 12, Application US/07626923A
; GENERAL INFORMATION:
; APPLICANT: Yoshimura, Akihiko
; APPLICANT: Longmore, Gregory D.
; APPLICANT: Lodish, Harvey
; TITLE OF INVENTION: MUTANT EPO RECEPTOR AND USES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
```



ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/626,923A  
FILING DATE: 13 December 1990  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WH190-08  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-626-923A-12

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 1 LAR 3

RESULT 11  
US-08-357-264-5  
Sequence 5, Application US/08357264  
Patent No. 5541077  
GENERAL INFORMATION:  
APPLICANT: BURNIE Mr., James P.  
APPLICANT: MATTHEWS Ms., Ruth C.  
TITLE OF INVENTION: FUNGAL STRESS PROTEINS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cushman, Darby & Cushman  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,264  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/152669  
FILING DATE: 16-NOV-1993  
APPLICATION NUMBER: US 663897  
FILING DATE: 14-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, Paul N.  
REGISTRATION NUMBER: 16773  
REFERENCE/DOCKET NUMBER: PNK/5544/202253/DJP

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-357-264-5

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 1 LSR 3

RESULT 12  
US-08-456-840-38  
Sequence 38, Application US/08456840  
Patent No. 5597908  
GENERAL INFORMATION:  
APPLICANT: Taddei-Peters, W. C.  
APPLICANT: Butler, Sandra M.  
TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo No. 5597908el  
STREET: 1330 Piccard Drive  
CITY: Rockville  
STATE: Maryland  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,840  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/266,407  
FILING DATE: 27-JUN-1994  
APPLICATION NUMBER: US 08/172,461  
FILING DATE: 21-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gormley, Mary E.  
REGISTRATION NUMBER: 34,409  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-258-5200  
TELEFAX: 301-977-0847  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-456-840-38

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 1 LSR 3

## RESULT 13

US-08-672-514-5  
; Sequence 5, Application US/08672514  
; Patent No. 5686248  
; GENERAL INFORMATION:  
; APPLICANT: BURNIE Mr., James P.  
; APPLICANT: MATTHEWS Ms., Ruth C.  
; TITLE OF INVENTION: FUNGAL STRESS PROTEINS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cushman, Darby & Cushman  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,514  
; FILING DATE: 28-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/152,669  
; FILING DATE: 16-NOV-1993  
; APPLICATION NUMBER: US 08/152669  
; FILING DATE: 16-NOV-1993  
; APPLICATION NUMBER: US 663897  
; FILING DATE: 14-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, Paul N.  
; REGISTRATION NUMBER: 16773  
; REFERENCE/DOCKET NUMBER: PNK/5544/202253/DJP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3000  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-514-5

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 1 LSR 3

## RESULT 14

US-07-789-184-173  
; Sequence 173, Application US/07789184  
; Patent No. 5688768  
; GENERAL INFORMATION:  
; APPLICANT: COUGHLIN, SHAUN R.  
; APPLICANT: SCARBOROUGH, ROBERT M.  
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND  
; TITLE OF INVENTION: RELATED PHARMACEUTICALS  
; NUMBER OF SEQUENCES: 223  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California

COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/789,184  
; FILING DATE: 19911107  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 22000-20502.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 34-0154  
; INFORMATION FOR SEQ ID NO: 173:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 5  
; OTHER INFORMATION: /note= "This position is R-NH2."  
US-07-789-184-173

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 3 LAR 5

## RESULT 15

US-08-406-192-12  
; Sequence 12, Application US/08406192  
; Patent No. 5739287  
; GENERAL INFORMATION:  
; APPLICANT: Wilbur, D. Scott  
; APPLICANT: Prathare, Pradip M  
; TITLE OF INVENTION: Biotinylated Cobalamins  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness  
; STREET: 1420 Fifth Avenue, Suite 2800  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: WA 98101-2333  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/406,192  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/224,831  
; FILING DATE: 08-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Renzoni, George E  
; REGISTRATION NUMBER: 37,919  
; REFERENCE/DOCKET NUMBER: RECL18947  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 682 8100  
; TELEFAX: (206) 224 0779  
; TELEX: 4938023  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-406-192-12

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 3 LAR 5

Search completed: December 8, 2003, 09:20:30  
Job time : 2.85859 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 3.55556 Seconds  
(without alignments)  
209.232 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11  
Sequence: 1 LXR 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	4	9 US-09-248-158-2	Sequence 2, Appli
2	9	81.8	4	10 US-09-818-656A-8	Sequence 8, Appli
3	9	81.8	4	11 US-09-852-910-5	Sequence 5, Appli
4	9	81.8	4	11 US-09-852-910-155	Sequence 155, App
5	9	81.8	4	12 US-10-357-467-19	Sequence 19, Appl
6	9	81.8	4	12 US-10-277-292-674	Sequence 674, App
7	9	81.8	4	12 US-10-280-340-674	Sequence 674, App
8	9	81.8	4	12 US-10-391-399-115	Sequence 115, App
9	9	81.8	4	12 US-10-391-399-116	Sequence 116, App
10	9	81.8	4	15 US-10-206-699-295	Sequence 295, App
11	9	81.8	4	15 US-10-028-075B-6	Sequence 6, Appli
12	9	81.8	4	15 US-10-023-206A-6	Sequence 6, Appli
13	9	81.8	4	16 US-10-244-709-6	Sequence 6, Appli
14	9	81.8	5	12 US-10-407-123-108	Sequence 108, App
15	9	81.8	5	12 US-10-302-817A-36	Sequence 36, Appl

16	9	81.8	5	14 US-10-006-557-7	Sequence 7, Appli
17	9	81.8	5	15 US-10-206-699-32	Sequence 32, Appl
18	9	81.8	6	10 US-09-486-734A-17	Sequence 17, Appl
19	9	81.8	6	10 US-09-486-734A-27	Sequence 27, Appl
20	9	81.8	6	10 US-09-770-102A-17	Sequence 17, Appl
21	9	81.8	6	11 US-09-866-512A-9	Sequence 9, Appli
22	9	81.8	6	11 US-09-500-700-73	Sequence 73, Appl
23	9	81.8	6	12 US-10-308-967-17	Sequence 17, Appl
24	9	81.8	6	12 US-09-373-182B-1	Sequence 1, Appli
25	9	81.8	6	12 US-10-131-616B-6	Sequence 6, Appli
26	9	81.8	6	12 US-10-277-292-693	Sequence 693, App
27	9	81.8	6	12 US-10-280-340-693	Sequence 693, App
28	9	81.8	6	15 US-10-163-198-80	Sequence 80, Appl
29	9	81.8	6	16 US-10-244-709-4	Sequence 4, Appli
30	9	81.8	7	9 US-09-248-158-1	Sequence 1, Appli
31	9	81.8	7	9 US-09-779-233-37	Sequence 37, Appl
32	9	81.8	7	9 US-09-779-233-38	Sequence 38, Appl
33	9	81.8	7	9 US-09-779-233-42	Sequence 42, Appl
34	9	81.8	7	9 US-09-779-233-43	Sequence 43, Appl
35	9	81.8	7	9 US-09-779-233-44	Sequence 44, Appl
36	9	81.8	7	9 US-09-796-264-26	Sequence 26, Appl
37	9	81.8	7	9 US-09-731-242A-42	Sequence 42, Appl
38	9	81.8	7	9 US-09-731-242A-43	Sequence 43, Appl
39	9	81.8	7	9 US-09-731-242A-45	Sequence 45, Appl
40	9	81.8	7	9 US-09-989-789-123	Sequence 123, App
41	9	81.8	7	9 US-09-989-789-124	Sequence 124, App
42	9	81.8	7	9 US-09-989-789-125	Sequence 125, App
43	9	81.8	7	9 US-09-989-789-126	Sequence 126, App
44	9	81.8	7	9 US-09-989-789-127	Sequence 127, App
45	9	81.8	7	9 US-09-989-789-129	Sequence 129, App

#### ALIGNMENTS

#### RESULT 1

US-09-248-158-2  
; Sequence 2, Application US/09248158  
; Patent No. US20020015678A1  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Zhengyu  
; TITLE OF INVENTION: Direct Adsorption Scintillation Assay  
; TITLE OF INVENTION: for Measuring Enzyme Activity and Assaying Biochemical  
; TITLE OF INVENTION: Processes  
; FILE REFERENCE: 342312000600  
; CURRENT APPLICATION NUMBER: US/09/248,158  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR FILING DATE: 60/074,272  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
; US-09-248-158-2

Query Match 81.8%; Score 9; DB 9; Length 4;  
Best Local Similarity 66.7%; Pred. No. 6.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 2 LAR 4

#### RESULT 2

US-09-818-656A-8  
; Sequence 8, Application US/09818656A  
; Patent No. US20020142381A1

```
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001191
; CURRENT APPLICATION NUMBER: US/09/818,656A
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-818-656A-8

Query Match      81.8%; Score 9; DB 10; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      2 LTR 4

RESULT 3
US-09-852-910-5
; Sequence 5, Application US/09852910
; Publication No. US20030096297A1
; GENERAL INFORMATION:
; APPLICANT: Hamm, Heidi
; APPLICANT: Gilchrist, Annette
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled Receptor S
; FILE REFERENCE: 2661-101
; CURRENT APPLICATION NUMBER: US/09/852,910
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/275,472
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammal
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: PAR-23
US-09-852-910-5

Query Match      81.8%; Score 9; DB 11; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      1 LSR 3

RESULT 4
US-09-852-910-155
; Sequence 155, Application US/09852910
; Publication No. US20030096297A1
; GENERAL INFORMATION:
; APPLICANT: Hamm, Heidi
; APPLICANT: Gilchrist, Annette
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled Receptor S
; FILE REFERENCE: 2661-101
; CURRENT APPLICATION NUMBER: US/09/852,910
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/275,472
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 271
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: G alpha t library linker sequence
US-09-852-910-155

Query Match      81.8%; Score 9; DB 11; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      2 LTR 4

RESULT 5
US-10-357-467-19
; Sequence 19, Application US/10357467
; Publication No. US20030194729A1
; GENERAL INFORMATION:
; APPLICANT: Abogadie, Fe C.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Walker, Craig
; APPLICANT: Colledge, Clark
; APPLICANT: Hillyard, David R.
; APPLICANT: Jimenez, Elsie
; TITLE OF INVENTION: Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, p.c.
; STREET: 1425 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/357,467
; FILING DATE: 04-Feb-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/142,080
; FILING DATE: 15-MAY-2000
; APPLICATION NUMBER: WO US97/12618
; FILING DATE: 21-JUL-1997
; APPLICATION NUMBER: US 08/684,742
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-256.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
```

```
;
; LOCATION: 4
; OTHER INFORMATION: /note="Xaa is
; gamma-carboxyglutamic acid"
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-357-467-19

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 LXR 3
      | |
Db      1 LAR 3

RESULT 6
US-10-277-292-674
; Sequence 674, Application US/10277292
; Publication No. US20030199470A1
; GENERAL INFORMATION:
; APPLICANT: FARIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: APAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOVITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/277,292
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 674
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-10-277-292-674

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 LXR 3
      | |
Db      2 LTR 4

RESULT 7
US-10-280-340-674
; Sequence 674, Application US/10280340
; Publication No. US20030207835A1
; GENERAL INFORMATION:
; APPLICANT: FARIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: APAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOVITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/280,340
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 674
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-10-277-292-674

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 LXR 3
      | |
Db      2 LTR 4

RESULT 8
US-10-391-399-115
; Sequence 115, Application US/10391399
; Publication No. US20030219806A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Lora, Jose M.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL 18607, 15603, 69318, 12303, 48000,
; TITLE OF INVENTION: 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921
; TITLE OF INVENTION: AND 33751 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0200NMIM
; CURRENT APPLICATION NUMBER: US/10/391,399
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/789,481
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 09/634,669
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 09/583,373
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 09/510,706
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 10/309,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/336,936
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 10/094,214
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US60/275,078
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; Remaining Prior Application data removed -.See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: N-glycosylation site
US-10-391-399-115

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LSR 4

RESULT 9
US-10-391-399-116
; Sequence 116, Application US/10391399
; Publication No. US20030219806A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Gluckmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Lora, Jose M.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: NOVEL 18607, 15603, 48000,
; TITLE OF INVENTION: 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921
; TITLE OF INVENTION: AND 33751 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0200NMIM
; CURRENT APPLICATION NUMBER: US/10/391,399
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/789,481
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 09/634,669
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 09/583,373
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 09/510,706
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 10/309,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/336,936
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 10/094,214
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US60/275,078
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N-glycosylation site
US-10-391-399-116

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LSR 4

RESULT 10
US-10-206-699-295
; Sequence 295, Application US/10206699
; Publication No. US20030100510A1

; GENERAL INFORMATION:
; APPLICANT: Sundaramoorthy, M.
; APPLICANT: Hudson, B.
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NCI Domain Hexamer
; FILE REFERENCE: MBHB 01-1017
; CURRENT APPLICATION NUMBER: US/10/206,699
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 295
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-206-699-295

Query Match      81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 1 LAR 3

RESULT 11
US-10-028-075B-6
; Sequence 6, Application US/10028075B
; Publication No. US20030113733A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Gene regulator
; FILE REFERENCE: 2183-5223US
; CURRENT APPLICATION NUMBER: US/10/028,075B
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EP 01203748.7
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: swiss/p36507/MPK2 Human
US-10-028-075B-6

Query Match      81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LAR 4

RESULT 12
US-10-029-206A-6
; Sequence 6, Application US/10029206A
; Publication No. US20030119720A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Oligopeptide treatment of anthrax
; FILE REFERENCE: 2183-5222US
```



; CURRENT APPLICATION NUMBER: US/10/029,206A  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/821,380  
; PRIOR FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: Patentin ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-10-029-206A-6

Query Match 81.8%; Score 9; DB 15; Length 4;  
Best Local Similarity 66.7%; Pred. No. 6.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 2 LAR 4

RESULT 13  
US-10-244-709-6  
; Sequence 6, Application US/10244709  
; Publication No. US20030129769A1  
; GENERAL INFORMATION:  
; APPLICANT: FRED E. REGNIER  
; TITLE OF INVENTION: AFFINITY SELECTED SIGNATURE PEPTIDES FOR PROTEIN IDENTIFICATION  
; FILE REFERENCE: 290.0001 0103  
; CURRENT APPLICATION NUMBER: US/10/244,709  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 60/203,227  
; PRIOR FILING DATE: 2000-05-05  
; PRIOR APPLICATION NUMBER: 60/208,184  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/208,372  
; PRIOR FILING DATE: 2000-05-31  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 6  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: C-terminal sequence  
US-10-244-709-6

Query Match 81.8%; Score 9; DB 16; Length 4;  
Best Local Similarity 66.7%; Pred. No. 6.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 2 LAR 4

RESULT 14  
US-10-407-123-108  
; Sequence 108, Application US/10407123  
; Publication No. US20030181381A1  
; GENERAL INFORMATION:  
; APPLICANT: Himmelspach, Michele  
; Schlokot, Uwe  
; Dörner, Friedrich  
; Fisch, Andreas  
; Eibl, Johann  
; TITLE OF INVENTION: Factor X Analogues With  
; a Modified Protease Cleavage Site  
; NUMBER OF SEQUENCES: 122  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/407,123  
; FILING DATE: 04-Apr-2003  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/367,791A  
; FILING DATE: 12-NO. US20030181381A1-1999  
; APPLICATION NUMBER: AT A 335/97  
; FILING DATE: 27-FEB-1997  
; APPLICATION NUMBER: WO PCT/AT98/00045  
; FILING DATE: 27-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ausenhus, Scott L.  
; REGISTRATION NUMBER: 42,471  
; REFERENCE/DOCKET NUMBER: 20695D-000700US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 108:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 108:  
US-10-407-123-108

Query Match 81.8%; Score 9; DB 12; Length 5;  
Best Local Similarity 66.7%; Pred. No. 6.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 2 LTR 4

RESULT 15  
US-10-302-817A-36  
; Sequence 36, Application US/10302817A  
; Publication No. US20030198978A1  
; GENERAL INFORMATION:  
; APPLICANT: ROZZELLE, James  
; APPLICANT: BOLCHAKOVA, Elena  
; TITLE OF INVENTION: THERMUS BROCKIANUS NUCLEIC ACID POLYMERASES  
; FILE REFERENCE: 4768US  
; CURRENT APPLICATION NUMBER: US/10/302,817A  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: 60/334,434  
; PRIOR FILING DATE: 2001-11-30  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 36  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Thermus brockianus  
US-10-302-817A-36

Query Match 81.8%; Score 9; DB 12; Length 5;  
Best Local Similarity 66.7%; Pred. No. 6.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
Db 1 LSR 3

Search completed: December 8, 2003, 09:34:04  
Job time : 4.55556 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 5.45455 Seconds  
(without alignments)  
116.399 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11

Sequence: 1 LXXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*

```
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	3	20	AAW84193
2	9	81.8	4	19	AAW48182
3	9	81.8	4	19	AAW49964
4	9	81.8	4	22	AAQ79019
5	9	81.8	4	23	AAE24457
6	9	81.8	4	23	AAE20561
7	9	81.8	4	23	AAE14417
8	9	81.8	4	23	AAW51953
9	9	81.8	4	23	AAW51957

10	9	81.8	4	24	ABU36682	G protein coupled
11	9	81.8	4	24	ABU36806	G protein coupled
12	9	81.8	4	24	ABU13761	Novel human transp
13	9	81.8	5	11	AAW03446	Accessory moiety d
14	9	81.8	5	15	AAW48999	Sequence of portio
15	9	81.8	5	15	AAW50135	Ovine growth hormo
16	9	81.8	5	15	AAW50136	Ovine growth hormo
17	9	81.8	5	15	AAW50134	Ovine growth hormo
18	9	81.8	5	16	AAW77333	Human apolipoprote
19	9	81.8	5	17	AAW03695	Cyclic pentapeptid
20	9	81.8	5	17	AAW03696	Cyclic pentapeptid
21	9	81.8	5	17	AAW00251	Subtilisin N62D/G1
22	9	81.8	5	17	AAW00252	Subtilisin N62D/G1
23	9	81.8	5	18	AAW37317	Mammalian haemoglo
24	9	81.8	5	21	AAW51458	AAV VP3 derived pe
25	9	81.8	5	22	AAE11109	Tryptic peptide #1
26	9	81.8	5	22	AAE11144	Chymotryptic pepti
27	9	81.8	5	22	AAE05461	MASP substrate #3
28	9	81.8	5	23	AAE26308	Human rBPI protein
29	9	81.8	5	23	ABG77841	Targetting peptide
30	9	81.8	5	23	AAW51952	PSA antibody prepa
31	9	81.8	5	23	AAW51956	PSA antibody prepa
32	9	81.8	5	24	ABP55341	Cucurbita sp. pept
33	9	81.8	5	24	ABP55342	Spinacia oleracea
34	9	81.8	5	24	ABP75051	Proteome analysis
35	9	81.8	6	15	AAW46808	Phytase derived pe
36	9	81.8	6	15	AAW50141	Ovine growth hormo
37	9	81.8	6	15	AAW50142	Ovine growth hormo
38	9	81.8	6	15	AAW50143	Ovine growth hormo
39	9	81.8	6	15	AAW55741	dsRNA-dependent ki
40	9	81.8	6	17	AAW90511	Hybridoma ATCC HB-
41	9	81.8	6	18	AAW74602	3-methylcholanthre
42	9	81.8	6	18	AAW33665	Complement antagon
43	9	81.8	6	18	AAW37324	Mammalian haemoglo
44	9	81.8	6	18	AAW26572	Peptide immunoreac
45	9	81.8	6	19	AAW76723	Pancreatic polypep

#### ALIGNMENTS

##### RESULT 1

AAW84193

ID AAW84193 standard; peptide; 3 AA.

XX AAW84193;

AC AAW84193;

XX 25-MAR-1999 (first entry)

DT Peptide comprising a proteinase site.

DE Proteinase site; bone morphogenetic fusion protein; bone binding site;

XX Proteinase site; bone morphogenetic fusion protein; bone binding site;

KW bone morphogenetic protein; transforming growth factor beta;

KW active fragment; wound healing; bone growth.

XX Unidentified.

OS Unidentified.

XX WO9855137-A1.

PN WO9855137-A1.

XX 10-DEC-1998.

PD 10-DEC-1998.

PF 02-JUN-1998; 98WO-US11189.

XX 02-JUN-1998; 98WO-US11189.

XX 03-JUN-1997; 97US-0868452.

PR 03-JUN-1997; 97US-0868452.

XX (HALL/) HALL F L.

PA (HANE/) HAN B.

PA (NIMN/) NIMNI M E.

PA (SHOR/) SHORS E C.

PA (WULL/) WU L.

XX Hall FL, Han B, Nimmi ME, Shors EC, Wu L;

PI Hall FL, Han B, Nimmi ME, Shors EC, Wu L;

DR WPI; 1999-059875/05.

XX New bone morphogenetic fusion proteins - comprising a purification

PT tag and a bone morphogenetic active fragment, used for enhancing

PT wound healing or bone growth

XX

PS Claim 8; Page 34; 64pp; English.

XX

CC The present peptide represents a proteinase site used in the creation

CC of the bone morphogenetic fusion proteins of the invention. The bone

CC morphogenetic fusion protein may contain some or all of the following

CC elements: a purification tag, a proteinase site, an ECM/bone binding

CC site, a second proteinase site, and a bone morphogenetic protein

CC active fragment. The fusion proteins of the invention also includes

CC proteins that have transforming growth factor beta active fragments

CC instead of bone morphogenetic protein active fragments. The bone

CC morphogenetic fusion proteins can be used for enhancing wound healing

CC or bone growth.

XX

SQ Sequence 3 AA;

Query Match 81.8%; Score 9; DB 20; Length 3;

Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 1 LTR 3

RESULT 2

ID AAW48182 standard; peptide; 4 AA.

AC AAW48182;

XX

DT 30-JUN-1998 (first entry)

XX

DE Conantokin peptide derivative.

XX

KW Conantokin; predatory cone snail; treatment; neurologic disorder;

KW psychiatric disorder; anticonvulsant; neuroprotective; analgesic.

KW HIV infection; ophthalmic indication; memory; learning defect;

KW cognitive defect.

XX

OS Synthetic.

XX

FT Key Location/Qualifiers

FT Modified-site 4 /note= "gamma-carboxyglutamic acid"

FT

FN WO9803541-A1.

XX

PD 29-JAN-1998.

XX

PF 21-JUL-1997; 97WO-US12618.

XX

PR 22-JUL-1996; 96US-0684742.

XX

XX (COGN-) COGNETIX INC.

PA (UTAH) UNIV UTAH RES FOUND.

XX

PI Abogadie FC, Colledge C, Cruz LJ, Hillyard DR, Jimenez E;

PI Layer RT, McCabe RT, Olivera BM, Rivier JE, Shen GS;

PI Walker C, Zhou L;

XX

DR WPI; 1998-120694/11.

XX

PT New conantokin peptide(s) - useful for e.g. treating neurologic or

PT psychiatric disorders, or the management of pain

XX

PS Claim 15; Page 98; 122pp; English.

XX

CC The present sequence is a conantokin peptide derivative, which can

CC be used to treat neurologic and psychiatric disorders, e.g. as an

CC anticonvulsant, neuroprotective or analgesic agent. Neurologic and

CC psychiatric disorders include epilepsy, convulsions, neurotoxic

CC injury (associated with conditions of hypoxia, anoxia or ischaemia,

CC which typically follow stroke, cerebrovascular accident, brain or

CC spinal cord trauma, myocardial infarct, physical trauma, drowning,

CC suffocation, perinatal asphyxia or hypoglycaemic events),

CC neurodegeneration (associated with Alzheimer's disease, senile

CC dementia, Amyotrophic Lateral Sclerosis, Multiple Sclerosis,

CC Parkinson's disease, Huntington's disease, Down's Syndrome,

CC Korsakoff's disease, schizophrenia, AIDS dementia, multi-infarct

CC dementia, Binswanger dementia and neuronal damage associated with

CC uncontrolled seizures), chemical toxicity (such as addiction, and

CC morphine, opiate, opioid and barbiturate tolerance), pain (acute,

CC chronic, migraine), anxiety, major depression, manic-depressive

CC illness, obsessive-compulsive disorder, schizophrenia and mood

CC disorders (such as bipolar disorder, unipolar depression, dysthymia

CC and seasonal affective disorder) and dystonia (movement disorder),

CC sleep disorder, muscle relaxation and urinary incontinence. The

CC peptide can also be used to treat HIV infection, ophthalmic

CC indication and memory, learning or cognitive defects.

XX

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 19; Length 4;

Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 1 LAR 3

RESULT 3

AAW49964

ID AAW49964 standard; peptide; 4 AA.

XX

AC AAW49964;

XX

DT 30-JUN-1998 (first entry)

XX

DE Conantokin peptide derivative.

XX

KW Conantokin; predatory cone snail; treatment; neurologic disorder;

KW psychiatric disorder; anticonvulsant; neuroprotective; analgesic.

KW HIV infection; ophthalmic indication; memory; learning defect;

KW cognitive defect.

XX

OS Synthetic.

XX

FT Key Location/Qualifiers

FT Modified-site 4 /note= "gamma-carboxyglutamic acid"

FT

FN WO9803189-A1.

XX

PD 29-JAN-1998.

XX

PF 21-JUL-1997; 97WO-US12652.

XX

PR 06-DEC-1996; 96US-0762377.

PR 22-JUL-1996; 96US-0684750.

XX

XX (COGN-) COGNETIX INC.

PA

XX Layer RT, McCabe RT, Zhou L;

PI

XX WPI; 1998-120469/11.

DR

XX

PT Use of conantokin peptide(s) - for treating disorders involving

PT excessive excitation of nerve cells by excitatory amino acids or

PT agonists of the N-methyl-D-aspartate receptor

XX Claim 27; Page 69; 122pp; English.

XX The present sequence is a conantokin peptide derivative, which can

CC be used to treat neurologic and psychiatric disorders, e.g. as an

CC anticonvulsant, neuroprotective or analgesic agent. Neurologic and

CC psychiatric disorders include epilepsy, convulsions, neurotoxic

CC injury associated with conditions of hypoxia, anoxia or ischaemia,

CC which typically follow stroke, cerebrovascular accident, brain or

CC spinal cord trauma, myocardial infarct, physical trauma, drowning,

CC suffocation, perinatal asphyxia or hypoglycaemic events),

CC neurodegeneration (associated with Alzheimer's disease, senile

CC dementia, Amyotrophic lateral Sclerosis, Multiple Sclerosis,

CC Parkinson's disease, Huntington's disease, Down's Syndrome,

CC Korsakoff's disease, schizophrenia, AIDS dementia, multi-infarct

CC dementia, Binswanger dementia and neuronal damage associated with

CC uncontrolled seizures), chemical toxicity (such as addiction, and

CC morphine, opiate, opioid and barbiturate tolerance), pain (acute,

CC chronic, migraine), anxiety, major depression, manic-depressive

CC illness, obsessive-compulsive disorder, schizophrenia and mood

CC disorders (such as bipolar disorder, unipolar depression, dysthymia

CC and seasonal affective disorder) and dystonia (movement disorder),

CC sleep disorder, muscle relaxation and urinary incontinence. The

CC peptide can also be used to treat HIV infection, ophthalmic

CC indication and memory, learning or cognitive defects.

XX Sequence 4 AA;

SQ Query Match 81.8%; Score 9; DB 19; Length 4;

Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 1 LAR 3

RESULT 4

ID AAG79019 standard; peptide; 4 AA.

XX AAG79019;

AC AAG79019;

XX 10-DEC-2001 (first entry)

XX Amino acid sequence of conantokin R domain III.

XX Conantokin; cone snail; nerve cell excitation; NMDA receptor; epilepsy;

KW N-methyl-D-aspartate receptor; pain; psychiatric disorder;

KW neurotoxic injury; hypoxia; anoxia; ischemia; neurodegeneration;

KW chemical toxicity; addiction; drug craving; psychiatric disorder;

KW anxiety; depression; obsessive compulsive disorder; schizophrenia;

KW mood disorder; ophthalmic disorder; neurological disorder; dystonia;

KW sleep disorder; muscle relaxation; urinary incontinence;

KW cognition enhancement; HIV infection.

XX Conus radiatus.

XX

XX Key Location/Qualifiers

FT Modified-site 4

FT /note= "gamma-carboxyglutamic acid"

XX

XX US6277825-B1.

XX

XX 21-AUG-2001.

XX

XX 20-JUL-1999; 99US-0357141.

XX

XX 22-JUL-1996; 96US-0684750.

PR 06-DEC-1996; 96US-0762377.

PR 21-JUL-1997; 97WO-US12652.

PR 10-FEB-1999; 99US-0142076.

PR 01-APR-1999; 99US-0283277.

XX (UTAH ) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

XX Olivera BM, McIntosh JM, McCabe RT, Layer RT, Zhou L;

PI WPI; 2001-601377/68.

XX

DR Use of conantokin peptide or its derivatives or a conantokin peptide

PT chimera for treating disorders e.g. migraine

PT Claim 9; Column 80; 60pp; English.

XX AAG79012-43 and AAG790054-56 represent domains of conantokin peptides.

CC Conantokin differ from conotoxins, in that they contain

CC gamma-carboxyglutamic acid. The conantokin are derived from the venom

CC of cone snails. They are used for the treatment of disorders in which

CC the pathophysiology involves excessive excitation of nerve cells by

CC excitatory amino acids or agonist of N-methyl-D-aspartate (NMDA)

CC receptor. The conantokin peptides are used for the treatment of

CC disorders such as pain; neurologic or psychiatric disorders such

CC epilepsy; for reducing neurotoxic injury associated with conditions of

CC hypoxia, anoxia or ischemia; for treating neurodegeneration; for

CC treating chemical toxicity such as addiction, drug craving, alcohol

CC abuse, morphine, opioid and barbiturate tolerance; for treating

CC psychiatric disorders such as anxiety, major depression, manic-depression

CC illness, obsessive compulsive disorder, schizophrenia or mood disorder;

CC for treating ophthalmic disorder; for treating additional neurological

CC disorders e.g. dystonia, sleep disorder, muscle relaxation and urinary

CC incontinence; for memory/cognition enhancement; for treating HIV

CC infection.

XX Sequence 4 AA;

SQ Query Match 81.8%; Score 9; DB 22; Length 4;

Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 1 LAR 3

RESULT 5

AAE24457

ID AAE24457 standard; peptide; 4 AA.

XX AAE24457;

AC AAE24457;

XX 04-OCT-2002 (first entry)

XX BoNT/A N-glycosylation site #1.

XX Neurotoxin; biological persistence; blepharospasm; pain; therapy;

KW neuromuscular disorder; cervical dystonia; oromandibular dystonia;

KW spasmodic dysphonia; laryngeal dystonia; muscular tension; asthma;

KW neuralgia; autonomic nervous system disorder; sweating; salivation;

KW headache; neuropathy; botulinum toxin serotype A; BoNT/A.

XX Clostridium botulinum.

XX WO200240506-A2.

XX

XX 23-MAY-2002.

XX

XX 16-NOV-2001; 2001WO-US44030.

XX

XX 17-NOV-2000; 2000US-249540P.

XX

XX (ALLR ) ALLERGAN SALES INC.

XX Steward LE, Spanoyannis A, Lin W, Aoki KR;

XX

DR WPI; 2002-479904/51.  
 XX Modified neurotoxin especially Clostridial toxins, useful for treating  
 PT neuromuscular and autonomic nervous system disorder and pain, comprises  
 PT structural modification to alter biological persistence of neurotoxin  
 PT -  
 XX  
 XX Disclosure; Page 19; 55pp; English.  
 XX  
 CC The invention related to modified neurotoxins especially  
 CC Clostridial botulinum toxins with altered biological persistence.  
 CC These toxins comprise a structural modification which is effective to  
 CC alter the biological persistence. Modified neurotoxins of the invention  
 CC are used for treating biological disorders which include neuromuscular  
 CC disorder e.g. strabismus, blepharospasm, spasmodic torticollis (cervical  
 CC dystonia), oromandibular dystonia and spasmodic dysphonia (laryngeal  
 CC dystonia), autonomic nervous system disorders e.g. excessive salivation  
 CC and sweating, asthma etc. and pain e.g. headache, muscular tension,  
 CC neuralgia and neuropathy. The present sequence is botulinum toxin  
 CC serotype A (BoNT/A) N-glycosylation site. This peptide is derived from  
 CC C. botulinum.  
 XX  
 SQ Sequence 4 AA;  
 Query Match 81.8%; Score 9; DB 23; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 DB 2 LTR 4  
 DE Soybean diverged delta-9 fatty acid desaturase peptide #7.  
 DE Soybean; diverged delta-9 fatty acid desaturase; enzyme; cellular lipid;  
 KW oil; transgenic plant; gene mapping; immunisation.  
 OS Glycine max.  
 XX WO200216565-A2.  
 XX 28-FEB-2002.  
 XX 22-AUG-2001; 2001WO-US26246.  
 XX 22-AUG-2000; 2000US-226996P.  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX Booth JR, Cahoon RE, Hitz WD, Kinney AJ, Yadav NS;  
 DR WPI; 2002-269353/31.  
 XX New delta-9 fatty acid desaturase polypeptides and polynucleotides,  
 PT useful in creating transgenic plants having altered levels of mono-,  
 PT poly- and unsaturated fatty acids and in increasing the unsaturation  
 PT levels in cellular lipids -  
 XX Claim 12; Page 48; 77pp; English.  
 XX  
 CC The present invention relates to diverged delta-9 fatty acid desaturase  
 CC proteins and polynucleotides encoding such proteins. The nucleic acid  
 CC sequences may be used to increase the level of unsaturation in cellular  
 CC lipids, including oil, in tissues when the enzyme is absent or rate-

CC limiting, to isolate cDNAs and genes encoding homologous proteins from  
 CC the same or other plant species and to create transgenic plants in which  
 CC the polypeptides are present at higher or lower levels than normal or in  
 CC cell types or developmental stages in which they are not normally found,  
 CC thus altering the level of mono-, poly- and unsaturated fatty acids in  
 CC those cells. They are useful as probes for genetic and physical gene  
 CC mapping and as markers, e.g. restriction fragment length polymorphism  
 CC (RFLP) markers. The peptides can be used to immunise animals to produce  
 CC antibodies specific for the peptides and proteins. The present sequence  
 CC is soybean diverged delta-9 fatty acid desaturase peptide (residues  
 CC 256-259).  
 XX  
 SQ Sequence 4 AA;  
 Query Match 81.8%; Score 9; DB 23; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 DB 2 LAR 4  
 DE C-terminal sequence of an arginine-containing peptide.  
 DE Protein quantification; protein identification; signature peptide;  
 KW mass spectrometric analysis; proteolytic peptide; affinity ligand;  
 KW arginine; isobaric peptide.  
 XX Unidentified.  
 XX WO200186306-A2.  
 XX 15-NOV-2001.  
 XX 04-MAY-2001; 2001WO-US14418.  
 XX 05-MAY-2000; 2000US-203227P.  
 XX 31-MAY-2000; 2000US-208184P.  
 XX 31-MAY-2000; 2000US-208372P.  
 XX (PURD ) PURDUE RES. FOUND.  
 XX (REGN/) REGNIER F E.  
 XX (CHAK/) CHAKRABORTY A B.  
 XX (DORM/) DORMADY S J.  
 XX (GENG/) G'ENG M.  
 XX (JIJJ/) JI J.  
 XX (RIGG/) RIGGS L D.  
 XX (SIOM/) SIOMA C S.  
 XX (WANG/) WANG S.  
 XX (ZHAN/) ZHANG X.  
 XX Regnier FE, Chakraborty AB, Dormady SJ, G'eng M, Ji J, Riggs LD;  
 PI Sioma CS, Wang S, Zhang X;  
 DR WPI; 2002-089810/12.  
 XX Analyzing differences in protein content in protein samples, useful for  
 PT identifying protein(s) in a sample or multiple proteins in a single  
 PT complex mixture, by employing mass spectrometric analysis of signature  
 PT peptides -  
 XX Example 9; Page 73; 106pp; English.  
 XX The invention relates to analysing differences in protein content among

CC plural protein samples comprising employing mass spectrometric analysis  
 CC of proteolytic peptide fragments. The method comprises  
 CC fragmenting at least a first protein sample and a second protein  
 CC sample to produce a first peptide pool and a second peptide pool;  
 CC isotopically labelling at least a portion of the peptides in at  
 CC least one of the pools so as to permit resolution of otherwise  
 CC identical peptides in the first and second peptide pools by mass  
 CC analysis; contacting peptides from at least a portion of both of the  
 CC peptide pools with a capture moiety to yield affinity-selected peptides  
 CC comprising an affinity ligand, where the capture moiety selects for the  
 CC affinity ligand; and analysing the affinity-selected peptides by mass  
 CC spectrometry to determine one or more differences between the first and  
 CC second samples. The method of the invention is useful for identifying one  
 CC or more proteins in a sample, and is typically useful for identifying  
 CC multiple proteins in a single complex mixture. The method is also useful  
 CC for quantifying proteins in a sample or several samples. The advantages  
 CC of the novel method are that it is easier to separate peptides than  
 CC proteins; native structure of the protein does not have to be maintained  
 CC during the analysis; structural variants do not interfere; and  
 CC putative proteins suggested from DNA database can be recognised by  
 CC using a signature peptide probe. The present method permits  
 CC identification of a protein in a complex sample without purifying the  
 CC protein or obtaining its composite peptide signature. The present  
 CC sequence is C-terminal fragment of a differentially acetylated  
 CC arginine-containing peptide which is used to generate fragment ions for  
 CC distinguishing isobaric peptides.

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 23; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 DB 2 LAR 4

RESULT 8

AAM51953  
 ID AAM51953 standard; peptide; 4 AA.

AC AAM51953;

DT 01-FEB-2002 (first entry)

DE PSA antibody preparation immunogen peptide #5.

XX PSA; prostate-specific antigen; prostatic carcinoma; pro-kallikrein 2;  
 KW human; antibody; immunogen.

OS Homo sapiens.

XX DE10032040-A1.

XX 29-NOV-2001.

XX 05-JUL-2000; 2000DE-1032040.

XX 24-MAY-2000; 2000DE-1025387.

XX (HOFF ) ROCHE DIAGNOSTICS GMBH.

XX WPI; 2002-042633/06.

XX Antibody comprising specificity for some truncated forms of  
 PT pro-prostate specific antigen, useful in immunoassays for differential  
 PT diagnosis of prostatic carcinoma -

PS Disclosure; Page 11; 16pp; German.

XX The present invention relates to an antibody specific for (-5,-6 and  
 CC -7)-proPSA (prostate-specific antigen), but with no reactivity with (-4)

CC or shorter forms of the protein. The antibody can be used to diagnose  
 CC prostatic carcinoma and to detect human pro-kallikrein 2. The present  
 CC sequence is a peptide which forms part of a peptide-containing immunogen  
 CC described in the exemplification of the invention.

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 23; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 DB 2 LSR 4

RESULT 9

AAM51957  
 ID AAM51957 standard; peptide; 4 AA.

AC AAM51957;

DT 01-FEB-2002 (first entry)

XX PSA antibody preparation immunogen peptide #9.

DE PSA; prostate-specific antigen; prostatic carcinoma; pro-kallikrein 2;  
 KW human; antibody; immunogen.

OS Homo sapiens.

XX DE10032040-A1.

XX 29-NOV-2001.

XX 05-JUL-2000; 2000DE-1032040.

XX 24-MAY-2000; 2000DE-1025387.

XX (HOFF ) ROCHE DIAGNOSTICS GMBH.

XX WPI; 2002-042633/06.

XX Antibody comprising specificity for some truncated forms of  
 PT pro-prostate specific antigen, useful in immunoassays for differential  
 PT diagnosis of prostatic carcinoma -

PS Disclosure; Page 11; 16pp; German.

XX The present invention relates to an antibody specific for (-5,-6 and  
 CC -7)-proPSA (prostate-specific antigen), but with no reactivity with (-4)  
 CC or shorter forms of the protein. The antibody can be used to diagnose  
 CC prostatic carcinoma and to detect human pro-kallikrein 2. The present  
 CC sequence is a peptide which forms part of a peptide-containing immunogen  
 CC described in the exemplification of the invention.

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 23; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 DB 1 LSR 3

RESULT 10

ABJ36682  
 ID ABJ36682 standard; Peptide; 4 AA.

XX ABJ36682;

XX









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 4.20202 Seconds  
(without alignments)  
245.646 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11  
Sequence: 1 LXR 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	8	Q16468	Q16468 homo sapien
2	9	81.8	8	Q9T4Y2	Q9T4Y2 asterina pe
3	9	81.8	8	O89965	O89965 polyomaviru
4	9	81.8	9	Q9T688	Q9T688 gecko gecko
5	9	81.8	9	Q94VH4	Q94VH4 varanus gla
6	9	81.8	9	Q94VD8	Q94VD8 varanus nil
7	9	81.8	9	Q94V18	Q94V18 varanus ere
8	9	81.8	9	Q94VC6	Q94VC6 varanus pil
9	9	81.8	9	Q94VE1	Q94VE1 varanus mer
10	9	81.8	9	Q91BM8	Q91BM8 simian viru
11	9	81.8	12	Q9PYK1	Q9PYK1 simian viru
12	9	81.8	10	Q958K9	Q958K9 rana boylli
13	9	81.8	10	Q9TG86	Q9TG86 diploglossu
14	9	81.8	10	Q94V97	Q94V97 varanus spe
15	9	81.8	10	Q94VD5	Q94VD5 varanus oli
16	9	81.8	10	Q94VC9	Q94VC9 varanus pan

17	9	81.8	10	8	Q94VF0	Q94VF0 varanus kin
18	9	81.8	10	8	Q9TG47	Q9TG47 ophisaurus
19	9	81.8	10	8	P92771	P92771 xenosaurus
20	9	81.8	10	8	Q94V85	Q94V85 varanus var
21	9	81.8	10	8	Q9TG32	Q9TG32 ophisaurus
22	9	81.8	10	8	Q94PD8	Q94PD8 varanus sca
23	9	81.8	10	8	Q9TG38	Q9TG38 ophisaurus
24	9	81.8	10	8	P92576	P92576 bipes bipor
25	9	81.8	10	8	Q9TG35	Q9TG35 ophisaurus
26	9	81.8	10	8	Q94VD2	Q94VD2 varanus pan
27	9	81.8	10	8	Q9TFV5	Q9TFV5 eublepharus
28	9	81.8	10	8	Q8SHN1	Q8SHN1 bradypodion
29	9	81.8	11	3	Q9C1R7	Q9C1R7 saccharomyc
30	9	81.8	11	7	Q97900	Q97900 oreochromis
31	9	81.8	11	7	Q97917	Q97917 oreochromis
32	9	81.8	11	7	Q97902	Q97902 oreochromis
33	9	81.8	11	7	Q97921	Q97921 pseudotroph
34	9	81.8	11	7	Q97901	Q97901 oreochromis
35	9	81.8	11	7	Q97916	Q97916 oreochromis
36	9	81.8	11	7	Q97905	Q97905 oreochromis
37	9	81.8	11	7	Q97899	Q97899 oreochromis
38	9	81.8	11	7	Q97812	Q97812 oreochromis
39	9	81.8	11	7	Q97904	Q97904 oreochromis
40	9	81.8	11	7	Q97903	Q97903 oreochromis
41	9	81.8	11	8	Q94V94	Q94V94 varanus sto
42	9	81.8	11	8	Q9G631	Q9G631 calotes nig
43	9	81.8	11	8	Q9G622	Q9G622 saalea horsf
44	9	81.8	11	8	Q94VG8	Q94VG8 varanus gou
45	9	81.8	11	8	Q9G365	Q9G365 calotes emm

#### ALIGNMENTS

##### RESULT 1

Q16468 PRELIMINARY; PRT; 8 AA.  
AC Q16468;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE DNA for cosmid CC13-1134 PCR primer 1 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96435920; PubMed=8838806;  
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,  
RA Anand R.;  
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5  
ESTs including 3 members of the Cystatin gene family and  
RT identification of CpG islands.";  
RL Genomics 32:425-430(1996).  
DR EMBL: X88976; CAA61407.1; -;  
FT NON\_TER  
SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871E6 CRC64;

Query Match 81.8%; Score 9; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
Db 5 LTR 7

##### RESULT 2

Q9T4Y2 PRELIMINARY; PRT; 8 AA.  
ID Q9T4Y2  
AC Q9T4Y2;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)



RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407514; AAL10096.1; -.  
KW Mitochondrion.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 3 LTR 5

RESULT 7  
Q94V18 ID Q94V18 PRELIMINARY; PRT; 9 AA.  
AC Q94V18;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Varanus eremias.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
OX NCBI\_TaxID=169838;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ast J.C.;  
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407495; AAL10040.1; -.  
KW Mitochondrion.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1124 MW; 9E80C733640DD731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 3 LAR 5

RESULT 8  
Q94VC6 ID Q94VC6 PRELIMINARY; PRT; 9 AA.  
AC Q94VC6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Varanus pilbarensis.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
OX NCBI\_TaxID=62048;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ast J.C.;  
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407518; AAL10108.1; -.  
KW Mitochondrion.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1064 MW; 874CASA36411A735 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 3 LTR 5

RESULT 9  
Q94VE1 ID Q94VE1 PRELIMINARY; PRT; 9 AA.  
AC Q94VE1;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Varanus mertensi.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
OX NCBI\_TaxID=62044;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ast J.C.;  
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407512; AAL10090.1; -.  
KW Mitochondrion.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 3 LTR 5

RESULT 10  
Q91BM8 ID Q91BM8 PRELIMINARY; PRT; 9 AA.  
AC Q91BM8;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Papovavirus BK (Gardner derived clone BKV9) early transcription control region (Fragment).  
OS Simian virus 12.  
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.  
OX NCBI\_TaxID=46771;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Gardner;  
RX MEDLINE=87061221; PubMed=3023684;  
RA Chuke W.F., Walker D.L., Peitzman L.B., Frisque R.J.;  
RT "Construction and characterization of hybrid polyomavirus genomes.";  
RL J. Virol. 60:960-971(1986).  
DR EMBL; M14452; AAA96236.1; -.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1130 MW; C7FD15B736C40732 CRC64;

Query Match 81.8%; Score 9; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 6 LSR 8

## RESULT 11

ID Q9PYK1 PRELIMINARY; PRT; 9 AA.  
AC Q9PYK1  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Papovavirus BK (Gardner) early transcription control region (Fragment).  
OS Simian virus 12.  
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.  
OX NCBI\_TaxID=46771;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Gardner;  
RX MEDLINE=87061221; PubMed=3023684;  
RA Chuks W.F., Walker D.L., Peitzman L.B., Frisque R.J.;  
RT "Construction and characterization of hybrid polyomavirus genomes.";  
RL J. Virol. 60:960-971(1986).  
DR EMBL; M14451; AAA96235.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1130 MW; C7FD15B736C40732 CRC64;

Query Match 81.8%; Score 9; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 6 LSR 8

## RESULT 12

ID Q958K9 PRELIMINARY; PRT; 10 AA.  
AC Q958K9  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Rana boylei.  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
OX NCBI\_TaxID=160499;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21184280; PubMed=1128498;  
RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,  
Jennings M., Larson A.;  
RT "Molecular Phylogenetics of Western North American Frogs of the Rana boylei Species Group."  
RL Mol. Phylogenet. Evol. 19:131-143(2001).  
DR EMBL; AF314019; AAK56877.1; -.  
KW Mitochondrion.  
FT NON TER 10  
SQ SEQUENCE 10 AA; 1320 MW; 42D380C9D36411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 3 LTR 5

## RESULT 13

ID Q9TG86 PRELIMINARY; PRT; 10 AA.  
AC Q9TG86;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Diploglossus bilobatus.  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidodactylus; Squamata; Scleroglossa; Anguimorpha; Anguillidae;  
OC Diploglossus.  
OX NCBI\_TaxID=102183;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99343613; PubMed=10413621;  
RA Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,  
Papenfuss T.J.;  
RT "Molecular phylogenetics, trna evolution, and historical biogeography in anguillid lizards and related taxonomic families."  
RL Mol. Phylogenet. Evol. 12:250-272(1999).  
DR EMBL; AF085608; AAD51514.1; -.  
KW Mitochondrion.  
FT NON TER 10  
SQ SEQUENCE 10 AA; 1255 MW; 5DEE80C7336411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 3 LTR 5

## RESULT 14

ID Q94V97 PRELIMINARY; PRT; 10 AA.  
AC Q94V97  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Varanus spenceri.  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidodactylus; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
OX NCBI\_TaxID=169854;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Med J.C.;  
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407530; AAL10142.1; -.  
KW Mitochondrion.  
FT NON TER 10  
SQ SEQUENCE 10 AA; 1255 MW; 5DEE80C7336411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 3 LTR 5

## RESULT 15

ID Q94VDS PRELIMINARY; PRT; 10 AA.  
AC Q94VDS  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Varanus olivaceus.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidossauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
OX NCBI\_TaxID=62047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Aet J.C.;  
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407515; AAL10099.1; -.  
KW Mitochondrion.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1234 MW; 584C4CA5A36411A7 CRC64;  
  
Query Match 81.8%; Score 9; DB 8; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 LXR 3  
Db 3 LTR 5

Search completed: December 8, 2003, 09:18:42  
Job time : 6.20202 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.05051 Seconds  
(without alignments)  
179.063 Million cell updates/sec

Title: US-09-498-556C-357  
Perfect score: 11  
Sequence: 1 LXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	%	ID	Description
1	9	81.8	11	RS30_ONCMY	P83328 oncorhynchu
2	9	81.8	15	ONC1_ONCMY	P83287 oncorhynchu
3	9	81.8	15	UC25_MAIZE	P80631 zea mays (m
4	9	81.8	16	IBP4_PIG	P24854 sus scrofa
5	9	81.8	18	VAA5_RHOPA	Q02005 rhodopseudo
6	9	81.8	20	LYC_FELCA	P37155 felis silve
7	9	81.8	20	RECA_AZOVI	P37863 azotobacter
8	9	81.8	21	SYA_RAT	P50475 rattus norv
9	9	81.8	22	ANFC_CHICK	P21805 gallus gall
10	9	81.8	23	GLNA_PHOLP	P20479 phormidium
11	9	81.8	25	IPYR_PSEAN	P80898 pseudanabae
12	9	81.8	25	NEUU_RANTE	P20056 rana tempor
13	9	81.8	25	SPIG_PSEUS	P23357 pseudacanth
14	9	81.8	25	UBLI_BOVIN	P23356 bos taurus
15	9	81.8	26	NTRC_RHILP	P41502 rhizobium 1
16	9	81.8	26	YFRA_KLEPN	P21710 klebsiella
17	9	81.8	27	CKX3_CONRA	P58806 conus radia
18	9	81.8	27	SECR_CANPA	P09910 canis famil
19	9	81.8	27	SECR_RABIT	P32647 corytolagus
20	9	81.8	27	SECR_SHEEP	P31299 ovis aries
21	9	81.8	28	VI03_VACCP	Q00334 vaccinia vi
22	9	81.8	29	HS98_NEUCR	P31540 neurospora
23	9	81.8	29	Y51_BPT3	P20326 bacterioph
24	9	81.8	31	A98A_DROME	O46201 drosophila
25	9	81.8	31	DIUX_DIPPU	P82372 diptoptera
26	9	81.8	33	LYC2_HORSE	P81710 equus cabal
27	9	81.8	33	OTCC_PSEPU	P11727 pseudomonas
28	9	81.8	33	PK1_DICDI	P34101 dictyosteli
29	9	81.8	34	RNL1_PIG	P15466 sus scrofa
30	9	81.8	36	F4RE_METOG	P80951 methanogeni
31	9	81.8	36	PAHO_CERSI	P37999 ceratotheri
32	9	81.8	36	PAHO_CHIBR	P41519 chinchilla
33	9	81.8	36	PAHO_DIDMA	P18107 didelphis m

34	9	81.8	36	1	PAHO_EQUZE	P38000 equus zebra
35	9	81.8	36	1	PAHO_ERIEU	P41335 erinaceus e
36	9	81.8	36	1	PAHO_MACMU	P33884 macaca mula
37	9	81.8	36	1	PAHO_RABIT	P41336 oryctolagus
38	9	81.8	36	1	PAHO_TAPPI	P39659 tapirus pin
39	9	81.8	36	1	PYU_AMICA	P29205 amia calva
40	9	81.8	36	1	PYU_PIG	P01305 sus scrofa
41	9	81.8	37	1	CALI_PIG	P30880 sus scrofa
42	9	81.8	37	1	CALI_SHEEP	P30881 ovis aries
43	9	81.8	37	1	CALR_RANRI	P31888 rana ridibu
44	9	81.8	37	1	PIF7_BOVIN	P21671 bos taurus
45	9	81.8	37	1	PSBY_GUITH	O78433 guillardia

ALIGNMENTS

RESULT 1  
RS30\_ONCMY STANDARD; PRT; 11 AA.  
AC P83328;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE 40S ribosomal protein S30 (Fragment).  
GN PAU.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]\_TaxID=8022;  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC TISSUE=Skin mucus;  
RX MEDLINE=22142142; PubMed=12147245;  
RA Fernandes J.M.O., Smith V.J.;  
RT "A novel antimicrobial function for a ribosomal peptide from rainbow trout skin";  
RL Biochem. Biophys. Res. Commun. 296:167-171(2002).  
CC -!- FUNCTION: Has antibacterial activity against Gram-positive bacteria.  
CC -!- MASS SPECTROMETRY: MW=6676.6; METHOD=MALDI.  
CC -!- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.  
KW Ribosomal protein; Antibiotic.  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1123 MW; 2312AB630DD735B8 CRC64;  
Query Match 81.8%; Score 9; DB 1; Length 11;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LXR 3  
Db 6 LAR 8

RESULT 2  
ONC1\_ONCMY STANDARD; PRT; 15 AA.  
AC P83287;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RC TISSUE=Skin;  
RA Fernandes J.M.O., Smith V.J., Kemp G.D.;

```

RT "Purification and N-terminal sequencing of a 3 kDa antibacterial
RL peptide from skin secretions of rainbow trout."
CC Submitted (MAY-2002) to the SWISS-PROT data bank.
CC -I- FUNCTION: Has antibacterial activity against Gram-positive
CC bacterium P.citreus.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Skin.
DR GO: GO:000576; C:extracellular; NAS.
DR GO: GO:0003795; F:antimicrobial peptide activity; NAS.
DR GO: GO:0006805; P: xenobiotic metabolism; NAS.
KW Antibiotic.
FT UNSURE 4 4 OR G.
FT NON TER 15 15 OR T.
FT SEQUENCE 15 AA; 1601 MW; 43C25028DBC12B7C CRC64;

Query Match 81.8%; Score 9; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 12 LAR 14

RESULT 3
UC25_MAIZE STANDARD; PRT; 15 AA.
AC P80631;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 39, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 77)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernellet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 4.9. ITS MW IS: 31.6 kDa.
DR Maize-2DPAGE; P80631; COLEOPTILE.
DR MaizeDB; 123957; -.
FT NON TER 1 1
FT NON TER 15 15
FT SEQUENCE 15 AA; 1580 MW; 83C54CF0CE1614D0 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 5 LSR 7

RESULT 4
IBP4_PIG STANDARD; PRT; 16 AA.
AC P24854;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein 4 (IGFBP-4) (IBP-4)
DE (IGF-binding protein 4) (Fragment).

```

---

```

GN IGFBP4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=92109718; PubMed=1722398;
RA Coleman M.E., Fan Y.-C.E., Ethernon T.D.;
RT "Identification and NH2-terminal amino acid sequence of three
RT insulin-like growth factor-binding proteins in porcine serum.";
RL Biochem. Biophys. Res. Commun. 181:1131-1136(1991).
CC -I- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: Contains 1 IGFBP domain.
CC -I- SIMILARITY: Contains 1 thyroglobulin type-I domain.
DR PIR; JH0517; JH0517.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR00716; Thyroglobulin_1.
DR PROSITE; PS00222; IGF BINDING; PARTIAL.
DR PROSITE; PS00484; THYROGLOBULIN_1; PARTIAL.
KW Growth factor binding
FT NON TER 16 16
FT SEQUENCE 16 AA; 1799 MW; 4098884009655E2 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 16;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 14 LAR 16

RESULT 5
YAA5_RHOPA STANDARD; PRT; 18 AA.
AC Q02005;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in aadR 5'region (fragment).
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RX MEDLINE=92394882; PubMed=1522059;
RA Dispensa M., Thomas C.T., Kim M.K., Perrotta J.A., Gibson J.,
RA Harwood C.S.;
RT "Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate
RT is dependent on AadR, a member of the cyclic AMP receptor protein
RT family of transcriptional regulators.";
RL J. Bacteriol. 174:5803-5813(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M92426; AAA26089.1; -.
DR PIR; A43334; A43334.
KW Hypothetical protein.
FT NON TER 1 1
FT SEQUENCE 18 AA; 2173 MW; 61DABDFE4E317B0A CRC64;

```

Query Match 81.8%; Score 9; DB 1; Length 18;  
Best Local Similarity 66.7%; Pred. No. 6.6e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXR 3  
DB 6 LTR 8

## RESULT 6

LYC\_FELCA STANDARD; PRT; 20 AA.  
AC P37155;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Lysozyme C (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase C) (Fragment).  
GN LYZ.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Milk;  
RX MEDLINE=90263403; PubMed=2344734;  
RA Halliday J.A., Bell K., McKenzie H.A., Shaw D.C.;  
RT "Feline whey proteins: identification, isolation and initial  
RT characterization of alpha-lactalbumin, beta-lactoglobulin and  
RT lysozyme";  
RL Comp. Biochem. Physiol. 95B:773-779(1990).  
CC -I- FUNCTION: Lysozymes have primarily a bacteriolytic function; those  
CC in tissues and body fluids are associated with the monocyte-  
CC macrophage system and enhance the activity of immunogens.  
CC -I- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-  
CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan  
CC heteropolymers of the prokaryotes cell walls.  
CC -I- SUBUNIT: Monomer.  
CC -I- MISCELLANEOUS: Lysozyme C is capable of both hydrolysis and  
CC transglycosylation; it shows also a slight esterase activity. It  
CC acts rapidly on both peptide-substituted and unsubstituted  
CC peptidoglycan, and slowly on chitin oligosaccharides.  
CC -I- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.  
CC PIR; A60525; A60525.  
DR HSP; P11376; 2EOL.  
DR InterPro; IPR001916; Glyco\_hydro\_22.  
DR Pfam; PF00062; Iys; 1.  
DR PROSITE; PS00128; LACTALBUMIN\_LYSOZYME; PARTIAL.  
KW Hydrolase; Glycosidase; Bacteriolytic enzyme; Milk.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2314 MW; EB8824EA59425E13 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 7.3e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXR 3  
DB 8 LAR 10

## RESULT 7

RECX\_AZOVI STANDARD; PRT; 20 AA.  
AC F37863;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Regulatory protein recX (Fragment).  
GN RECX.

OS Azotobacter vinelandii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Azotobacter.  
OX NCBI\_TaxID=354;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92225347; PubMed=1563632;  
RA Venkatesh T.V., Das H.K.;  
RT "The Azotobacter vinelandii recA gene: sequence analysis and  
RT regulation of expression";  
RL Gene 113:47-53(1992).  
RN [2]  
RN IDENTIFICATION.  
RX MEDLINE=94218258; PubMed=8165147;  
RA de Mot R., Schoofs G., Vanderleyden J.;  
RT "A putative regulatory gene downstream of recA is conserved in gram-  
RT negative and Gram-positive bacteria";  
RL Nucleic Acids Res. 22:1313-1314(1994).  
CC -I- FUNCTION: Modulates recA activity (By similarity).  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -I- SIMILARITY: BELONGS TO THE RECX FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; S96898; -; NOT ANNOTATED\_CDS.  
DR HAMAP; MF\_01114; -; 1.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2111 MW; C809F8BCED6CB56 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 7.3e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXR 3  
DB 18 LAR 20

## RESULT 8

SYA\_RAT STANDARD; PRT; 21 AA.  
ID SYA\_RAT  
AC P50475;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Alanine-tRNA synthetase (EC 6.1.1.7) (Alanine-tRNA ligase) (AlaRS)  
DE (Fragment).  
GN AARS.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.

RC TISSUE=Liver;  
RX MEDLINE=91249799; PubMed=2040280;  
RA Dignam J.D., Dignam S.S., Brumley L.L.;  
RT "Alanine-tRNA synthetase from Escherichia coli, Bombyx mori and Ratus  
RT ratus. Existence of common structural features";  
RL Eur. J. Biochem. 198:201-210(1991).  
CC -I- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +  
CC diphosphate + L-alanyl-tRNA(Ala).  
CC -I- SUBUNIT: Monomer.  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -I- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
CC PIR; S16073; S16073.  
DR InterPro; IPR006193; tRNA synt Ala.  
DR PROSITE; PS50860; AA TRNA\_LIGASE II ALA; PARTIAL.  
KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding.

```

FT NON TER 1 1
FT NON TER 21 21
SQ SEQUENCE 21 AA; 2293 MW; D739DDC62CD43375 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 7.7e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 18 LAR 20

RESULT 9
ANFC_CHICK
ID ANFC_CHICK STANDARD; PRT; 22 AA.
AC P21805;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-type natriuretic peptide (CNP).
GN NPFC.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91113186; PubMed=1989595;
RA Arimura J.J., Minamino N., Kangawa K., Matsuo H.;
RT "Isolation and identification of C-type natriuretic peptide in
RT chicken brain.";
RL Biochem. Biophys. Res. Commun. 174:142-148(1991).
CC -|- FUNCTION: VASORELAXANT ACTIVITY. HAS A CGMP-STIMULATING ACTIVITY.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.
DR PIR; JT0581; JT0581.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP.1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
KW Vasoactive.
FT DISULFID 6 22
SQ SEQUENCE 22 AA; 2244 MW; FC2A4706D8DAC025 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LSR 4

RESULT 10
GLNA_PHOLP
ID GLNA_PHOLP STANDARD; PRT; 23 AA.
AC P20479;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase)
DE (Fragment).
OS Phormidium lapideum.
OC Bacteria; Cyanobacteria; Oscillatoriales; Phormidum.
OX NCBI_TaxID=32060;
RN [1]
RP SEQUENCE.
RX MEDLINE=89214011; PubMed=2907514;
RA Sawa Y., Ochiai H., Yoshida K., Tanizawa K., Tanaka H., Soda K.;

RT "Glutamine synthetase from a cyanobacterium, Phormidium lapideum:
RT purification, characterization, and comparison with other
RT cyanobacterial enzymes.";
RL J. Biochem. 104:917-923(1988).
CC -|- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -|- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC HEXAGONS.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR PIR; PX0011; PX0011.
DR InterPro; IPR001691; GLN_synth.
DR PROSITE; PS00180; GLNA_1; PARTIAL.
DR PROSITE; PS00181; GLNA_ATP; PARTIAL.
KW Ligase.
FT NON TER 23 23
SQ SEQUENCE 23 AA; 2656 MW; 20B69C164D2A5739 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 23;
Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 7 LSR 9

RESULT 11
IPYR_PSEAN
ID IPYR_PSEAN STANDARD; PRT; 25 AA.
AC P80898;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (PPase) (Fragment).
GN PPA.
OS Pseudanabaena sp. (strain PCC 6901).
OC Bacteria; Cyanobacteria; Oscillatoriales; Pseudanabaena.
OX NCBI_TaxID=47918;
RN [1]
RP SEQUENCE.
RA Gomez R., Serrano A.;
RL Submitted (FEB-1997) to the SWISS-PROT data bank.
CC -|- FUNCTION: Hydrolyzes PPI generated in anabolic reactions.
CC -|- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC -|- COFACTOR: Requires the presence of divalent metal cation.
CC Magnesium confers the highest activity. Binds 4 divalent cations
CC per subunit (By similarity).
CC -|- SUBUNIT: Homohexamer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the PPase family.
DR HAMAP; MF_00209; -; 1.
KW Hydrolase; Magnesium.
FT NON TER 25 25
SQ SEQUENCE 25 AA; 2630 MW; FB6138F3AED4D43E CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LSR 5

RESULT 12
NEUU_RANTE
ID NEUU_RANTE STANDARD; PRT; 25 AA.
AC P20056;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Neuromedin U-25 (NmU-25).  
OS Rana temporaria (European common frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
OX NCBI\_TaxID=8407;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Intestine;  
RX MEDLINE=90078173; PubMed=2592357;  
RA Domin J., Yiangou Y.G., Spokes R.A., Aitken A., Parmar K.B.,  
RA Chrysanthou B.J., Bloom S.R.;  
RT "The distribution, purification, and pharmacological action of an  
amphibian neuromedin U";  
RL J. Biol. Chem. 264:20881-20885(1989).  
CC -!- FUNCTION: STIMULATES UTERINE SMOOTH MUSCLE CONTRACTION AND CAUSES  
SELECTIVE VASOCONSTRICTION.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE NMU FAMILY.  
DE PIR; A34179; A34179.  
DR InterPro; IPR001942; NMU.  
DR PFam; PF02070; NMU; 1.  
DR SMART; SM00084; NMU; 1.  
DR PROSITE; PS00967; NMU; 1.  
KW Amidation; Hormone.  
FT MOD RES 25 25  
SQ SEQUENCE 25 AA; 2832 MW; 6A01D89F6DA06FD4 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;  
Best Local Similarity 66.7%; Pred. No. 9.3e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 14 LSR 16

RESULT 13  
ID SPIG\_PSEUS STANDARD; PRT; 25 AA.  
AC P82357;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Spingerin.  
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Isoptera; Termitidae;  
OC Macrotermitinae; Pseudacanthotermes.  
OX NCBI\_TaxID=115113;  
RN [1]  
RP SEQUENCE, MASS SPECTROMETRY, AND FUNCTION.  
RC TISSUE=Blood, and Salivary gland;  
RX PubMed=11053427;  
RA Lamberty M., Zachary D., Lanot R., Bordereau C., Robert A.,  
RA Hoffmann J.A., Bulet P.;  
RT "Insect Immunity. Constitutive expression of a cysteine-rich  
antifungal and a linear antibacterial peptide in a termite insect.";  
RL J. Biol. Chem. 276:4085-4092(2001).  
CC -!- FUNCTION: ACTIVE AGAINST GRAM-POSITIVE BACTERIA B.MEGATERIUM AND  
M.LUTEUS, GRAM-NEGATIVE BACTERIA E.COLI SBS363 AND D22,  
K.PNEUMONIAE, S.TYPHIMURIUM AND P.AERUGINOSA, YEAST C.ALBICANS  
AND FILAMENTOUS FUNGI F.CULMORUM, N.CRASSA, N.HEMATOCOCCA AND  
T.VIRIDAE. INACTIVE AGAINST GRAM-POSITIVE BACTERIA B.SUBTILIS,  
S.PYOGENES, B.THURINGIENSIS AND S.AUREUS, GRAM-NEGATIVE BACTERIA  
E.CLOACAE AND E.CAROTOFORA AND FILAMENTOUS FUNGUS B.BASSIANA.  
CC -!- INDUCTION: By bacterial infection.  
CC -!- MASS SPECTROMETRY: MW=3001.8; METHOD=WALDI.  
CC -!- MISCELLANEOUS: THERE ARE THREE ISOFORMS OF SPINGERIN.  
KW Antibiotic; Fungicide.  
FT VARIANT 1 3 MISSING (IN N-3 ISOFORM).  
FT VARIANT 22 25 MISSING (IN C-4 ISOFORM).  
SQ SEQUENCE 25 AA; 3001 MW; AA79370264262F60 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;  
Best Local Similarity 66.7%; Pred. No. 9.3e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 22 LTR 24

RESULT 14  
ID UBLI\_BOVIN STANDARD; PRT; 25 AA.  
AC P23356;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-  
L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)  
DE (PGP 9.5) (PGP9.5) (Fragment).  
GN UCHL1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=92008646; PubMed=1833240;  
RA Giambanco I., Bianchi R., Ceccarelli P., Pula G., Sorci G.,  
RA Antonoli S., Bocchini V., Donato R.;  
RT "Neuron-specific" protein gene product 9.5 (PGP 9.5) is also  
expressed in glioma cell lines and its expression depends on cellular  
growth state.";  
RL FEBS Lett. 290:131-134(1991).  
CC -!- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE  
PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUINATED PROTEINS.  
CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE  
A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.  
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =  
ubiquitin + a thiol.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- TISSUE SPECIFICITY: NEURONS AND CELLS OF THE DIFFUSE  
NEUROENDOCRINE SYSTEM AND THEIR TUMORS.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.  
DR PIR; S17561; S17561.  
DR MEROPS; C12.001; -.  
DR InterPro; IPR001578; UCH 1.  
DR PFam; PF01088; Peptidase\_C12; 1.  
DR PROSITE; PS00140; UCH 1; PARTIAL.  
KW Ubl conjugation pathway; Hydrolase; Thiol protease; Multigene family.  
FT NON TER 25 25  
SQ SEQUENCE 25 AA; 2812 MW; 26BB5ADD0A754D55 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;  
Best Local Similarity 66.7%; Pred. No. 9.3e+02;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 17 LTR 19

RESULT 15  
ID NTRC\_RHILP STANDARD; PRT; 26 AA.  
AC P41502;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Nitrogen assimilation regulatory protein (Fragment).  
GN NTRC.  
OS Rhizobium leguminosarum (biovar phaseoli).

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=385;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CE-3;
RX MEDLINE=94018651; PubMed=8412703;
RA Patriarca E.J., Riccio A., Tate R., Colonna-Romano S., Iaccarino M.,
RA Defez R.;
RT "The ntrBC genes of Rhizobium leguminosarum are part of a complex
RT operon subject to negative regulation.";
RL Mol. Microbiol. 9:569-577(1993).
CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM NTRB/NTRC
CC INVOLVED IN THE ACTIVATION OF NITROGEN ASSIMILATORY GENES SUCH AS
CC GUNA. NTRC IS PHOSPHORYLATED BY NTRB AND INTERACTS WITH SIGMA-54.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X71436; -; NOT_ANNOTATED_CDS.
DR PIR; S36203;
DR InterPro; IPR001789; Response reg.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
KW Nitrogen fixation; Transcription regulation; Repressor; Activator;
KW DNA-binding; ATP-binding; Phosphorylation; Sensory transduction.
FT DOMAIN 1 >26 RESPONSE REGULATORY.
FT NON_TER 26
FT SEQUENCE 26 AA; 2687 MW; E11ECCG3C8G304 CRC64;
SQ
Query Match 81.8%; Score 9; DB 1; Length 26;
Best Local Similarity 66.7%; Pred. No. 9.7e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 LXR 3
Db 22 LSR 24

```

Search completed: December 8, 2003, 09:14:15  
Job time : 2.05051 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.65657 Seconds  
(without alignments)  
232.212 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11

Sequence: 1 LXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 76:\*

2: pir1:\*

3: pir2:\*

4: pir3:\*

5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	7	E48394	glycoprotein compo
2	9	81.8	7	B48394	major fat-globule
3	9	81.8	8	PC4131	hypothetical prote
4	9	81.8	8	T13818	cytochrome oxidase
5	9	81.8	10	T13838	cytochrome-c oxida
6	9	81.8	10	T14219	cytochrome-c oxida
7	9	81.8	11	PU0034	dextranucrase (EC
8	9	81.8	12	A42324	cytochrome P450c27
9	9	81.8	13	PT0293	19 heavy chain CRD
10	9	81.8	14	PA0111	protein QA100054 -
11	9	81.8	15	S36896	ribosomal protein
12	9	81.8	15	PA0075	fructose-bisphosph
13	9	81.8	15	PA0102	fructose-bisphosph
14	9	81.8	15	B32800	hypothetical prote
15	9	81.8	15	PH1455	phe leader peptide
16	9	81.8	15	AF0832	T-cell receptor al
17	9	81.8	16	C61414	chymotrypsin (EC 3
18	9	81.8	16	JH0517	insulin-like growt
19	9	81.8	16	B44036	collagen alpha 1(X
20	9	81.8	17	JP0046	ribosomal protein
21	9	81.8	17	A34835	ribosomal protein
22	9	81.8	17	I54269	vitamin D binding
23	9	81.8	17	JQ2310	hypothetical 2.1K
24	9	81.8	17	JQ2320	hypothetical 2.1K
25	9	81.8	17	B61414	chymotrypsin (EC 3
26	9	81.8	18	S29491	GTP-binding protei
27	9	81.8	18	A41877	LcrKc - Yersinia p
28	9	81.8	18	A43334	orf1 5' of aadR -
29	9	81.8	18	S33645	hypothetical prote

30	9	81.8	18	2	S10452	hypothetical prote
31	9	81.8	19	2	S02269	glycogen(starch) s
32	9	81.8	19	2	UC2060	homeobox 5 protein
33	9	81.8	19	2	A60326	cholecystokinin-58
34	9	81.8	20	2	A60525	lysozyme (EC 3.2.1
35	9	81.8	20	2	JP0050	ribosomal protein
36	9	81.8	20	2	JP0051	ribosomal protein
37	9	81.8	20	2	S33001	hypothetical prote
38	9	81.8	20	2	PC2084	serine proteinase
39	9	81.8	20	2	S19616	globin - polychaet
40	9	81.8	20	2	A61414	chymotrypsin (EC 3
41	9	81.8	20	2	A42267	J-kappa recombinat
42	9	81.8	20	2	AC0269	probable trp opero
43	9	81.8	21	2	S16073	alanine-tRNA ligas
44	9	81.8	22	2	JT0581	natriuretic peptid
45	9	81.8	22	2	F23734	insulin-like growt

## ALIGNMENTS

### RESULT 1

E48394

glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fr  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997

C:Accession: E48394

R:Mather, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig

II-like sequences.

A:Reference number: A48394; MUID:93250576; PMID:8485470

A:Accession: E48394

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-7 <MAT>

A:Experimental source: milk

A>Note: sequence extracted from NCBI backbone (NCBIP:131450)

C:Keywords: glycoprotein

Query Match 81.8%; Score 9; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3

Db 5 LAR 7

### RESULT 2

B48394

major fat-globule membrane protein GP 55 - guinea pig (fragment)

C:Species: Cavia porcellus (guinea pig)

C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995

C:Accession: B48394

R:Mather, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig

II-like sequences.

A:Reference number: A48394; MUID:93250576; PMID:8485470

A:Accession: B48394

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-7 <MAT>

A:Experimental source: milk

A>Note: sequence extracted from NCBI backbone (NCBIP:131444)

Query Match 81.8%; Score 9; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3

Db 5 LAR 7

Db 5 LAR 7

RESULT 3  
PC4131  
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)  
C;Species: Pseudomonas aeruginosa  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C;Accession: PC4131  
R;Kawabaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.  
Gene 167, 87-91, 1995  
A;Title: Sequencing and characterization of the downstream region of the genes encoding Y for biosynthesis of heme d1.  
A;Reference number: JC4552; MUID:96144254; PMID:8566817  
A;Accession: PC4131  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-8 <RAW>  
A;Cross-references: DDBJ:D50473; NID:g1217594  
A;Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match 81.8%; Score 9; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 4 LSR 6

RESULT 4  
T13818  
Cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)  
C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: T13818  
R;Delarbre, C.; Barriuel, V.; Tillier, S.; Janvier, P.; Gachelin, G.  
Mol. Biol. Evol. 14, 807-813, 1997  
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI  
A;Reference number: Z17775; MUID:97398704; PMID:9254918  
A;Accession: T13818  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-8 <DEL>  
A;Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022  
C;Genetics:  
A;Genome: mitochondrion  
A;Note: COI  
C;Keywords: mitochondrion

Query Match 81.8%; Score 9; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 3 LSR 5

RESULT 5  
T13838  
cytochrome-c oxidase (EC 1.9.3.1) chain I - Bipes biporus mitochondrion (fragment)  
C;Species: mitochondrion Bipes biporus  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000  
C;Accession: T13838  
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.  
Mol. Biol. Evol. 14, 91-104, 1997  
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement  
A;Reference number: Z17789; MUID:97153826; PMID:9000757  
A;Accession: T13838  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-10 <MAC>

A;Cross-references: EMBL:U71335; NID:g1753232; PID:g1753235; PIDN:AAB48271.1  
C;Genetics:  
A;Genome: mitochondrion  
A;Note: COI  
C;Keywords: mitochondrion; oxidoreductase

Query Match 81.8%; Score 9; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 3 LTR 5

RESULT 6  
T14219  
cytochrome-c oxidase (EC 1.9.3.1) chain I - Xenosaurus grandis mitochondrion (fragment)  
C;Species: mitochondrion Xenosaurus grandis  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: T14219  
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.  
Mol. Biol. Evol. 14, 91-104, 1997  
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement  
A;Reference number: Z17789; MUID:97153826; PMID:9000757  
A;Accession: T14219  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-10 <MAC>  
A;Cross-references: EMBL:U71333; NID:g5739536; PIDN:AAC62821.1; PID:g1753275  
C;Genetics:  
A;Genome: mitochondrion  
A;Note: COI  
C;Keywords: mitochondrion; oxidoreductase

Query Match 81.8%; Score 9; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 3 LTR 5

RESULT 7  
PU0034  
dextranucrase (EC 2.4.1.5) - Streptococcus bovis (fragment)  
C;Species: Streptococcus bovis  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 18-Sep-1996  
C;Accession: PU0034  
R;Uezono, Y.; Tsumori, H.; Mukasa, H.  
submitted to JIPID, October 1993  
A;Description: Purification and properties of glucosyltransferase synthesizing 1,6-alpha-  
A;Reference number: PU0034  
A;Accession: PU0034  
A;Molecule type: protein  
A;Residues: 1-11 <UEZ>  
A;Experimental source: ATCC 9809  
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 81.8%; Score 9; DB 2; Length 11;  
Best Local Similarity 66.7%; Pred. No. 1.3e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
| |  
Db 8 LTR 10

RESULT 8  
A42324  
cytochrome P450c27/25 - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)



C;Date: 04-Mar-1993 #sequence\_revision 19-Nov-1994 #text\_change 08-Feb-1996  
C;Accession: A42324  
R;Shayiq, R.M.; Avadhani, N.G.  
J. Biol. Chem. 267, 2421-2428, 1992  
A;Title: Sequence complementarity between the 5'-terminal regions of mRNAs for rat mitochondria  
A;Reference number: A42324; MUID:92129322; PMID:1733943  
A;Accession: A42324  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-12 <SHA>  
A;Note: sequence extracted from NCBI backbone (NCBIN:78404, NCBIIP:78410)

Query Match 81.8%; Score 9; DB 2; Length 12;  
Best Local Similarity 66.7%; Pred. No. 1.4e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 4 LSR 6  
|

RESULT 9  
PT0293  
Ig heavy chain CDR3 region (clone 4-124) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C;Accession: PT0293  
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J<sub>H</sub> gene segments in the development of B lymphocyte  
A;Reference number: PT0222; MUID:91108337; PMID:1899102  
A;Accession: PT0293  
A;Molecule type: DNA  
A;Residues: 1-13 <YAM>  
A;Experimental source: B lymphocyte  
C;Keywords: heterotrimer; immunoglobulin

Query Match 81.8%; Score 9; DB 2; Length 13;  
Best Local Similarity 66.7%; Pred. No. 1.5e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 1 LSR 3  
|

RESULT 10  
PA0111  
protein OA100054 - Arabidopsis thaliana (fragment)  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 07-Apr-1995 #sequence\_revision 26-May-1995 #text\_change 06-Jun-1997  
C;Accession: PA0111  
R;Kamo, M.; Kawakami, T.; Tsugita, A.  
submitted to JIPID, March 1995  
A;Reference number: PA0109  
A;Accession: PA0111  
A;Molecule type: protein  
A;Residues: 1-14 <KAM>  
A;Experimental source: root

Query Match 81.8%; Score 9; DB 2; Length 14;  
Best Local Similarity 66.7%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 11 LSR 13  
|

RESULT 11  
S36896  
ribosomal protein S16 - Mycobacterium bovis (fragment)

C;Species: Mycobacterium bovis  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
C;Accession: S36896  
R;Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.  
FEBS Lett. 331, 9-14, 1993  
A;Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobacterium bovis  
A;Reference number: S36887; MUID:94009653; PMID:8405418  
A;Accession: S36896  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-15 <OHA>  
C;Keywords: protein biosynthesis; ribosome

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 6 LTR 8  
|

RESULT 12  
PA0075  
fructose-bisphosphate aldolase (EC 4.1.2.13) I - fungus (Fusarium sporotrichioides) (fragment)  
N;Alternate names: aldolase; fructose-1,6-bisphosphate triosephosphate-lase  
C;Species: Fusarium sporotrichioides  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
C;Accession: PA0075; PA0077  
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.  
submitted to JIPID, October 1994  
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides  
A;Reference number: PA0051  
A;Accession: PA0075  
A;Molecule type: protein  
A;Residues: 1-15 <CHO>  
A;Note: this form (II) had a molecular weight of 30.6K and an isoelectric point of 5.3  
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 6 LSR 8  
|

RESULT 13  
PA0102  
fructose-bisphosphate aldolase (EC 4.1.2.13) III - fungus (Fusarium sporotrichioides) (fragment)  
C;Species: Fusarium sporotrichioides  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
C;Accession: PA0102  
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.  
submitted to JIPID, October 1994  
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides  
A;Reference number: PA0051  
A;Accession: PA0102  
A;Molecule type: protein  
A;Residues: 1-15 <CHO>  
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 6 LSR 8  
|

Db 6 LSR 8

## RESULT 14

B32800  
hypothetical protein (P1 5' region) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 30-Sep-1993  
C:Accession: B32800  
R:Jindal, S.; Dudani, A.K.; Singh, B.; Harley, C.B.; Gupta, R.S.  
Mol. Cell. Biol. 9, 2279-2283, 1989  
A:Title: Primary structure of a human mitochondrial protein homologous to the bacterial  
A:Reference number: A32800; MUID:89313783; PMID:2568584  
A:Accession: B32800  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-15 <JIN>  
A:Cross-references: GB:M22382

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
Db 3 LSR 5

## RESULT 15

PH1455  
T-cell receptor alpha chain (clone A24/PEF4) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 11-Apr-1995  
C:Accession: PH1455  
R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; K  
J. Exp. Med. 177, 811-820, 1993  
A:Title: T cell receptor selection by and recognition of two class I major histocompatib  
A:Reference number: PH1430; MUID:93171821; PMID:8436911  
A:Accession: PH1455  
A:Molecule type: mRNA  
A:Residues: 1-15 <CAS>  
A:Experimental source: cytolytic T-lymphocyte  
C:Superfamily: immunoglobulin homology  
C:Keywords: receptor; T-cell

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
Db 3 LSR 5

Search completed: December 8, 2003, 09:19:34  
Job time : 2.65657 secs